

# Vorlesung 2013-2014: Lineare Algebra I und II

Hermann Schulz-Baldes<sup>\*†</sup>

4. Mai 2016

## Inhaltsverzeichnis

<b>1</b>	<b>Gruppen, Ringe, Körper</b>	<b>2</b>
<b>2</b>	<b>Die Struktur des Vektorraums</b>	<b>12</b>
<b>3</b>	<b>Lineare Gleichungssysteme und der Gauß-Algorithmus</b>	<b>21</b>
<b>4</b>	<b>Lineare Abbildungen und ihre darstellenden Matrizen</b>	<b>32</b>
<b>5</b>	<b>Permutationen und die Determinante</b>	<b>47</b>
<b>6</b>	<b>Eigenwerte und Eigenvektoren</b>	<b>57</b>
<b>7</b>	<b>Vektorräume mit Skalarprodukt</b>	<b>67</b>
<b>8</b>	<b>Jordan'sche Normalform</b>	<b>82</b>
<b>9</b>	<b>Matrixzerlegungen</b>	<b>94</b>
<b>10</b>	<b>Sesquilinearformen, hermitische Formen und Quadriken</b>	<b>109</b>
<b>11</b>	<b>Matrixnormen</b>	<b>122</b>
<b>12</b>	<b>Lineare Optimierung</b>	<b>131</b>
<b>A</b>	<b>Äquivalenzrelationen</b>	<b>140</b>
<b>B</b>	<b>Konvexitätsungleichungen</b>	<b>141</b>
	<b>Literatur</b>	<b>142</b>

---

<sup>\*</sup>Department Mathematik, Universität Erlangen-Nürnberg, Cauerstr. 11, D-91058 Erlangen, Germany

<sup>†</sup>e-mail: schuba@mi.uni-erlangen.de, web: www.mi.uni-erlangen.de/~schuba

**Voraussetzungen** (Rudimente aus der Schule bekannt):

- Elementare Logik (Teil des Vorkurses, ggfs. lesen Sie Kapitel 3 von Schichl, Steinbauer)
- Begriff der natürlichen, ganzen und reellen Zahlen (Details in Analysis)
- Elementare Mengenlehre (Details in Analysis)
- Begriff der Abbildung zwischen Mengen (Details in Analysis)

# 1 Gruppen, Ringe, Körper

**1.1 Definition** Sei  $G$  eine Menge versehen mit Abbildung (binärer Operation)  $\circ : G \times G \rightarrow G$ . Dann ist  $(G, \circ, e)$  eine Gruppe  $\iff$  (i), (ii), (iii) gelten, wobei

- (i)  $\forall a, b, c \in G$  gilt  $a \circ (b \circ c) = (a \circ b) \circ c$  (Assoziativität)
- (ii)  $\exists$  neutrales Element  $e \in G$  mit  $e \circ a = a$
- (iii)  $\forall a \in G \exists b \in G$  mit  $b \circ a = e$  (Existenz von Inversen)  
Schreibweise:  $b = a^{-1}$

Falls lediglich (i) und (ii) gelten, heißt  $(G, \circ, e)$  Halbgruppe.

Falls zudem  $a \circ b = b \circ a \quad \forall a, b \in G$ , heißt  $(G, \circ, e)$  kommutativ oder abelsch.

Eine Teilmenge  $H \subset G$  heißt Untergruppe  $\iff (H, \circ, e)$  Gruppe

**1.2 Beispiele** 1. Die einfachste nicht-triviale Gruppe  $\mathbb{Z}_2$  besteht aus zwei Elementen  $\mathbb{Z}_2 = \{0, 1\}$ . Das neutrale Element ist  $e = 0$  und die Verknüpfungstafel ist

$\circ$	0	1
0	0	1
1	1	0

2. Sei  $\mathbb{N} = \{1, 2, 3, \dots\}$  die Menge der natürlichen Zahlen. Hierauf gibt es die Operation *Addition zweier Zahlen*, bezeichnet mit  $+$ . Dennoch ist  $\mathbb{N}$  keine Gruppe, sogar keine Halbgruppe, weil es kein neutrales Element gibt.
3. Sei  $\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$  wieder versehen mit der Addition. Dann ist  $(\mathbb{N}_0, +, 0)$  eine Halbgruppe mit neutralem Element  $e = 0$ . Es ist aber keine Gruppe, da keine Inverse existieren.
4. Sei  $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$  die Menge der ganzen Zahlen, auf der wieder die Addition  $+$  definiert ist. Dann ist  $(\mathbb{Z}, +, 0)$  eine kommutative Gruppe, wobei  $e = 0$  und Inverses  $(a)^{-1} = -a$ .
5. Auch die rationalen Zahlen  $\mathbb{Q} = \{\frac{p}{q} \mid p \in \mathbb{Z}, q \in \mathbb{N}\}$  und die reellen Zahlen  $\mathbb{R}$  (die Zahlengerade, Konstruktion siehe Analysis) versehen mit der Addition bilden kommutative Gruppen  $(\mathbb{Q}, +, 0)$  und  $(\mathbb{R}, +, 0)$ . Zudem ist  $(\mathbb{Z}, +, 0)$  Untergruppe von sowohl  $(\mathbb{Q}, +, 0)$  wie auch  $(\mathbb{R}, +, 0)$ , und  $(\mathbb{Q}, +, 0)$  Untergruppen von  $(\mathbb{R}, +, 0)$  sind.
6. Sei  $p \in \mathbb{N}$ . Die Menge  $p\mathbb{Z} = \{pm \mid m \in \mathbb{Z}\}$  ist eine Untergruppe von  $(\mathbb{Z}, +, 0)$ , und somit auch von  $(\mathbb{R}, +, 0)$ , und  $(\mathbb{Q}, +, 0)$

7. Die Menge  $\mathbb{R} \setminus \{0\}$  versehen mit der Multiplikation ist eine Gruppe. Eine Untergruppe davon ist  $\mathbb{Q} \setminus \{0\}$ .

8. Sei  $X$  eine Menge und  $A = \{f \mid f : X \rightarrow X \text{ Abbildung}\}$ .

Auf  $A$  gibt es eine binäre Operation  $\circ : A \times A \rightarrow A$  gegeben durch die Hintereinanderausführung.

Dann ist  $(A, \circ, \text{id}_X)$  eine Halbgruppe mit  $e = \text{id}_X$ . Andererseits existiert im Allgemeinen kein Inverses. Um dies sicherzustellen, müssen wir die Untermenge  $G \subset A$  von invertierbaren Funktionen betrachten. Eine Funktion heißt invertierbar (oder bijektiv oder eineindeutig), wenn Bilder unterschiedlicher Elemente unterschiedlich sind (Injektivität) und jedes Element Bildpunkt ist (Surjektivität).

Weiter ist  $G = \{f \in A \mid f \text{ bijektiv}\}$  eine Gruppe (Inverses existiert!), die im Allgemeinen nichtkommutativ ist.

Falls  $X = \{1, 2, \dots, n\}$ , so heißt  $S_n = \{f \mid f = \{1, \dots, n\} \rightarrow \{1, \dots, n\} \text{ bijektiv}\}$  die symmetrische Gruppe, über die wir noch viel mehr erfahren werden.

9. Gruppen treten oft als Symmetriegruppen auf. Als Beispiel betrachten wir ein regelmäßiges  $n$ -Eck. Dem zugeordnet sind:

Diedergruppe  $D_{2n} = \{\text{Rotationen um } k\frac{2\pi}{n}, k = 0, \dots, n-1 \text{ um den Ursprung, Spiegelungen}\}$

Untergruppe  $\mathbb{Z}_n = \{\text{Rotationen um } k\frac{2\pi}{n}, k = 0, \dots, n-1 \text{ um den Ursprung}\} \subset D_{2n} \quad \diamond$

**1.3 Bemerkung** Wie aus den Beispielen hervorgeht, ist die Verknüpfung mit  $\circ$ ,  $+$  oder  $\cdot$  bezeichnet. Der Kürze wegen wird oft auch einfach nichts geschrieben, d.h.  $ab$  für  $a \circ b$ , insbesondere wenn keine Verwechslung möglich ist.  $\diamond$

**1.4 Satz** Sei  $G$  eine Gruppe mit Eins  $e$  und seien  $a, b, c \in G$ . Dann gilt:

- (i)  $a^{-1}a = e \implies aa^{-1} = e$  (Linksinverse = Rechtsinverse)
- (ii)  $ea = a \iff ae = a$
- (iii) Es gibt nur ein neutrales Element  $e$ .
- (iv) Auch das Inverse ist eindeutig.
- (v)  $\exists! x \in G$  mit  $ax = b$
- (vi)  $\exists! x \in G$  mit  $xa = b$
- (vii) Kürzungsregel:  $ab = ac \implies b = c$
- (viii)  $(a^{-1})^{-1} = a$
- (ix)  $(ab)^{-1} = b^{-1}a^{-1}$

### Begründung

(i) Sei  $a^{-1}a = e$ , dann  $\exists b \in G$  mit  $ba^{-1} = e$ . Unter Verwendung der Assoziativität,

$$\begin{aligned} aa^{-1} &= e(aa^{-1}) = (ba^{-1})(aa^{-1}) = b(a^{-1}a)a^{-1} \\ &= b(ea^{-1}) = ba^{-1} = e. \end{aligned}$$

(ii) Gegeben  $ea = a$  und  $a^{-1}a = e$

$$ae = aa^{-1}a \stackrel{(i)}{=} ea = a .$$

(iii) Gegeben zwei Einsen  $e, e'$ , so gilt

$$e = e'e \stackrel{(ii)}{=} e' .$$

(iv) Seien  $a^{-1}, b$  zwei Inverse von  $a$ . Dann

$$b = be = baa^{-1} = ea^{-1} = a^{-1} .$$

(v) Setze  $x = a^{-1}b$ . Dann gilt  $ax = aa^{-1}b = b$ . Sei  $x'$  eine zweite Lösung  $ax' = b$ . Dann

$$x' = a^{-1}ax' = a^{-1}b = a^{-1}ax = x .$$

(vi) und (vii) Analog bzw. Übung

(viii)

$$(a^{-1})^{-1} = (a^{-1})^{-1}e = (a^{-1})^{-1}a^{-1}a = ea = a .$$

(ix)

$$(ab)^{-1} = (ab)^{-1}aa^{-1} = (ab)^{-1}abb^{-1}a^{-1} = b^{-1}a^{-1} .$$

□

**1.5 Definition** Seien  $(G_1, e_1)$  und  $(G_2, e_2)$  Gruppen. Eine Abbildung  $h : G_1 \rightarrow G_2$  heißt ein Homomorphismus genau dann, wenn

$$h(ab) = h(a)h(b) \quad \forall a, b \in G_1$$

Wenn  $h$  zudem bijektiv (eindeutig) ist, heißt  $h$  ein Isomorphismus. Die beiden Gruppen heißen dann isomorph.

**1.6 Bemerkungen** 1. Beachten Sie, dass in der Gleichung  $h(ab) = h(a)h(b)$  links die Verknüpfung in  $G_1$ , und rechts in  $G_2$  verwandt wird.

2. Es gilt  $h(e_1) = h(e_1e_1) = h(e_1)h(e_1)$ , also  $e_2 = h(e_1)$  nach der Eindeutigkeit des Einselementes.

3. Anstelle von Homomorphismus spricht man auch von einer *strukturhaltenden Abbildung*, und die Struktur ist hier die Gruppenstruktur gegeben durch die Verknüpfung. Es gibt auch andere Strukturen, die erhalten werden können, und deswegen ist es manchmal hilfreich zu spezifizieren und im jetzigen Fall von einem Gruppenhomomorphismus zu sprechen. Analoges gilt für den Isomorphismus.

4. Isomorphe Gruppen können von einem gruppentheoretischen Standpunkt aus nicht unterschieden werden und deswegen einfach als ein und dieselbe Gruppe aufgefaßt werden. ◇

**1.7 Beispiel** 1. Sei  $n \in \mathbb{N}$ . Die Abbildung  $h : \mathbb{Z} \rightarrow \mathbb{Z}$  gegeben durch  $h(m) = nm$  ist ein Homomorphismus, denn  $h(m+k) = n(m+k) = nm+nk = h(m) + h(k)$ . Das Bild  $h(\mathbb{Z}) = n\mathbb{Z}$  ist eine Untergruppe von  $\mathbb{Z}$ . Fasst man  $h$  auf als  $\hat{h} : \mathbb{Z} \rightarrow n\mathbb{Z}$ , so ist  $\hat{h}$  ein Isomorphismus (Führen Sie den expliziten Beweis durch).

2. Die Abbildung  $h : \mathbb{Z} \rightarrow \mathbb{Z}_2$  gegeben durch  $h(m) = m \bmod 2 = \begin{cases} 0, & m \text{ gerade} \\ 1, & m \text{ ungerade} \end{cases}$  ist ein Gruppenhomomorphismus. ◇

**1.8 Satz** Sei  $h : G_1 \rightarrow G_2$  ein Homomorphismus zwischen zwei Gruppen  $(G_1, e_1)$  und  $(G_2, e_2)$ . Dann gilt:

- (i) Der Kern von  $h$  definiert als  $\text{Ker}(h) = \{a \in G_1 \mid h(a) = e_2\}$  ist eine Untergruppe von  $G_1$ .
- (ii) Das Bild von  $h$  definiert als  $\text{Ran}(h) = \{b \in G_2 \mid \exists a \in G_1 : h(a) = b\}$  ist eine Untergruppe von  $G_2$ .

**Beweis.** (i) Wenn  $a, b \in \text{Ker}(h)$ , dann  $h(ab) = h(a)h(b) = e_2e_2 = e_2$ , so dass  $ab \in \text{Ker}(h)$ . Somit führt die Verknüpfung nicht aus  $\text{Ker}(h)$  heraus. Außerdem ist das neutrale Element  $e_1 \in \text{Ker}(h)$  und, für  $a \in \text{Ker}(h)$  gilt  $e_1 = h(aa^{-1}) = h(a)h(a^{-1}) = e_2h(a^{-1}) = h(a^{-1})$ , so dass auch das Inverse  $a^{-1} \in \text{Ker}(h)$ . Teil (ii) ist eine Übung.  $\square$

Nun führen wir ein weiteres gruppentheoretisches Konzept ein.

**1.9 Definition** Sei  $G$  eine Gruppe mit Untergruppe  $H$  und  $a \in G$ . Dann sind die Rechtsnebenklassen  $Ha$  und Linksnebenklassen  $aH$  definiert durch

$$Ha = \{xa \mid x \in H\}, \quad aH = \{ax \mid x \in H\}.$$

Die Untergruppe  $H$  heißt Normalteiler von  $G$ , falls  $Ha = aH$  für alle  $a \in G$ . Die Menge aller Linksnebenklassen wird mit  $G/H$  bezeichnet.

**1.10 Bemerkungen**

1. Weil  $e \in H$ , gilt  $a \in aH$  und  $a \in Ha$ .
2.  $aH = bH$  ist äquivalent zu  $b^{-1}a \in H$ . In der Tat,  $aH = bH$  impliziert  $a = ae \in bH$ , d.h.  $a = bh$  für ein  $h \in H$ . Umgekehrt,  $b^{-1}a \in H$  impliziert  $b^{-1}a = h$  für ein  $h \in H$ , so dass  $a = bh \in bH$ , was zu  $aH \subset bH$  führt. Die umgekehrte Inklusion folgt aus  $a^{-1}b = (b^{-1}a)^{-1} \in H$ .
3. Es gilt entweder  $aH = bH$  oder  $aH \cap bH = \emptyset$ .
4. Der Kern eines Gruppenhomomorphismus  $h : G_1 \rightarrow G_2$  ist ein Normalteiler von  $G_1$ . Des Weiteren ist  $G_1/\text{Ker}(h)$  isomorph zu  $\text{Ran}(h)$ . All dies zu verifizieren ist eine Übung.
5. Die Rechtsnebenklassen sind genau die Äquivalenzklassen der Äquivalenzrelation  $x \sim y \iff \exists b \in G : xb \in H$  und  $yb \in H$ . Sobald dieser Begriff aus der Analysis bekannt ist (siehe auch den Appendix), überzeugen Sie sich davon im Detail. Dies impliziert auch, dass die Zugehörigkeit zu Links- bzw. Rechtsnebenklassen eine Äquivalenzrelation ist (was auch direkt verifiziert werden kann). Deswegen wird oft auch  $[a]$  anstelle von  $Ha$  geschrieben.  $\diamond$

**1.11 Beispiel** Sei  $p \in \mathbb{N}$  und betrachte die Untergruppe  $p\mathbb{Z} \subset \mathbb{Z}$ . Die Nebenklassen werden mit  $[n] = n(p\mathbb{Z})$  bezeichnet und heißen auch Restklassen. Konkret sind sie gegeben durch

$$[n] = \{n + kp \mid k \in \mathbb{Z}\} = n + p\mathbb{Z}.$$

Es gibt also  $p$  unterschiedlich Nebenklassen  $[0], \dots, [p-1]$ . Offensichtlich sind die Rechts- und Linksnebenklassen gleich. Somit ist  $p\mathbb{Z}$  ein Normalteiler von  $\mathbb{Z}$ . Die Menge der Nebenklassen wird bezeichnet mit

$$\mathbb{Z}_p = \{[n] \mid n \in \mathbb{Z}\} = \{[0], [1], \dots, [p-1]\}.$$

Auf  $\mathbb{Z}_p$  wird nun eine Addition  $+$  :  $\mathbb{Z}_p \times \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  definiert durch  $[n] + [m] = [n + m]$ . In der Tat ist die rechte Seite unabhängig von der Wahl der Repräsentanten und zudem ist die Addition kommutativ. Außerdem ist  $[0]$  ein neutrales Element und ein Inverses ist gegeben durch  $[n]^{-1} = [-n]$ . Zusammenfassend ist also  $(\mathbb{Z}_p, +, [0])$  kommutative Gruppe, genannt die zyklische Gruppe von Ordnung  $p$ . (Die Ordnung einer endlichen Gruppe ist immer die Anzahl ihrer Elemente, und die ist hier  $p$ .) Dieses Beispiel wird nun verallgemeinert.  $\diamond$

**1.12 Satz** Sei  $H$  ein Normalteiler einer Gruppe  $G$ . Durch  $aH \circ bH = abH$  wird  $G/H$  eine Gruppe mit neutralem Element  $eH$ , die Faktorgruppe oder Quotientengruppe heißt.

**Beweis** Zunächst muss überprüft werden, dass die Verknüpfung  $aH \circ bH$  tatsächlich wohldefiniert, d.h. unabhängig von der Wahl der Repräsentanten  $aH$  und  $bH$  der Nebenklassen ist. Sei also  $aH = a'H$  und  $bH = b'H$ . Dann ist  $a^{-1}a' \in H$  und  $b^{-1}b' \in H$ . Somit ist

$$(ab)^{-1}a'b' = b^{-1}(a^{-1}a')b' \in b^{-1}Hb' = b^{-1}b'H = H, ,$$

wobei im vorletzten Schritt die Voraussetzung verwendet wurde. Also ist  $(ab)H = (a'b')H$ , d.h.  $aH \circ bH = a'H \circ b'H$ . Nun kann die Assoziativität überprüft werden:

$$(aH \circ bH) \circ cH = abH \circ cH = abcH = a(bc)H = aH \circ (bc)H = aH \circ (bH \circ cH) .$$

Das neutrale Element von  $G/H$  ist  $eH$  weil  $eH \circ aH = eaH = aH$ . Zu guter Letzt ist das inverse Element gegeben durch  $(aH)^{-1} = a^{-1}H$ .  $\square$

Nun werden Mengen betrachtet, die neben der Gruppenstruktur noch eine weitere binäre Operation besitzen.

**1.13 Definition** Sei  $R$  eine Menge mit zwei Operatoren  $+, \cdot : R \times R \rightarrow R$

Dann ist  $(R, +, \cdot, 0)$  ein Ring  $\iff$

- (i)  $(R, +, 0)$  ist eine kommutative Gruppe mit neutralem Element 0
- (ii) Das Assoziativgesetz für  $\cdot$  gilt:  $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
- (iii) Die Distributivgesetze gelten:  $(a + b) \cdot c = a \cdot c + b \cdot c$  und  $a \cdot (b + c) = a \cdot b + a \cdot c$

Zudem:

- (a)  $(R, +, \cdot, 0)$  heißt kommutativ  $\iff a \cdot b = b \cdot a$
- (b)  $(R, +, \cdot, 0)$  ist ein Ring mit Eins 1  $\iff \exists 1 \in R$  mit  $1 \cdot a = a \cdot 1 = a \quad \forall a \in R$ .

**1.14 Bemerkung** Das (additive) Inverse in  $(R, +, 0)$  wird meist mit  $-$  bezeichnet, d.h.  $a - a = 0$ . Der Punkt  $\cdot$  wird oft weggelassen.  $\diamond$

**1.15 Beispiele** 1.  $(\mathbb{Z}, +, \cdot, 0)$  und  $(\mathbb{Q}, +, \cdot, 0)$  sind kommutative Ringe mit Eins.

2.  $(\mathbb{Z}_p, +, \cdot, [0])$  ist kommutativer Ring mit Eins  $[1]$ , wobei die Multiplikation definiert ist durch  $[n] \cdot [m] = [nm]$ , was unabhängig von der Wahl der Repräsentanten ist (Übung). Oft wird  $\mathbb{Z}_p$  auch als Restklassenring bezeichnet.

3. Die Menge der Polynome in der Variablen  $x$  und mit Koeffizienten in den reellen Zahlen  $\mathbb{R}$  ist definiert als

$$\mathbb{R}[x] = \left\{ p(x) = \sum_{k=0}^n p_k x^k \mid n \in \mathbb{N}_0, p_k \in \mathbb{R}, p_n \neq 0 \right\} \cup \{0\} .$$

Hierbei ist das Summenzeichen definiert durch

$$\sum_{k=0}^n p_k x^k = p_0 x^0 + p_1 x^1 + p_2 x^2 + \dots + p_n x^n .$$

Dann heißt  $k$  der Summationsindex. Beachten Sie, dass die rechte Seite und somit auch die linke unabhängig von  $k$  ist und er somit  $k$  durch jedes andere Zeichen ersetzt werden kann. Auf der Menge der Polynome ist die Gradabbildung  $\text{Grad} : \mathbb{R}[x] \rightarrow \mathbb{N}_0 \cup \{-\infty\}$  definiert als

$$\text{Grad}(p) = \begin{cases} n & , \text{ falls } p(x) = \sum_{k=0}^n p_k x^k \text{ mit } p_n \neq 0 \\ -\infty & , \text{ falls } p = 0 \end{cases} .$$

Für  $p, q \in \mathbb{R}[x]$ , sind dann  $p + q \in \mathbb{R}[x]$  und  $p \cdot q \in \mathbb{R}[x]$  definiert durch

$$(p + q)(x) = p(x) + q(x) = \sum_{k=0}^n (p_k + q_k) x^k, \quad n = \max\{\text{Grad}(p), \text{Grad}(q)\}$$

$$(p \cdot q)(x) = p(x)q(x) = \sum_{k=0}^n \left( \sum_{\ell=0}^k p_\ell \cdot q_{k-\ell} \right) x^k, \quad n = \text{Grad}(p) + \text{Grad}(q) .$$

Hierbei werden alle fehlenden Koeffizienten durch 0 ergänzt. Mit diesen Operationen wird dann  $(\mathbb{R}[x], +, \cdot, 0)$  zu einem kommutativen Ring mit Eins  $e(x) = 1$ .

4. Eines der zentralen Objekte der linearen Algebra ist der Ring der quadratischen Matrizen versehen mit der Matrizenaddition und Matrizenmultiplikation. All dies wird später eingeführt.  $\diamond$

**1.16 Satz** Sei  $(R, +, \cdot, 0)$  ein Ring. Dann gilt für alle  $a, b \in R$ :

- (i)  $a \cdot 0 = 0 \cdot a = 0$
- (ii)  $-(ab) = (-a)b = a(-b)$
- (iii)  $ab = (-a)(-b)$
- (iv) Falls  $R$  eine Eins 1 hat, gilt zudem  $(-1) \cdot a = -a = a(-1)$

**Beweis.**

- (i)  $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0 \implies a \cdot 0 = 0$  (nach der Kürzungsregel in der Gruppe  $(R, +, 0)$ )
- (ii)  $ab + (-a)b = (a + (-a))b = 0b = 0$ . Somit  $(-a)b = -(ab)$ . Analog wird  $a(-b) = -(ab)$  gezeigt.
- (iii) + (iv) Übung  $\square$

**1.17 Definition**  $(K, +, \cdot, 0, 1)$  heißt Körper  $\iff (K, +, \cdot, 0)$  ist ein Ring mit Null 0 und Eins 1, so dass  $0 \neq 1$  und  $(K \setminus \{0\}, \cdot, 1)$  ist eine abelsche Gruppe.

**1.18 Bemerkungen** 1. Anstelle von  $a^{-1}b$  in  $(K \setminus \{0\}, \cdot, 1)$  wird auch die Schreibweise  $\frac{1}{a}b = \frac{b}{a}$  verwendet.

2. Es gilt folgende Umformulierung:  $(K, +, \cdot, 0, 1)$  Körper  $\iff$

- (i)  $(K, +, 0)$  abelsche Gruppe
- (ii)  $(K \setminus \{0\}, \cdot, 1)$  abelsche Gruppe
- (iii) Distributivgesetz  $(a + b)c = ac + bc$
- (iv)  $0 \neq 1$

◇

**1.19 Beispiele** 1.  $\mathbb{Q}$  und  $\mathbb{R}$  sind Körper, aber  $\mathbb{Z}$  nicht.

2. Sei  $\mathbb{Z}_2 = \{[0], [1]\} = \{0, 1\}$ . Auf  $\mathbb{Z}_2$  werden zwei Operationen durch folgende Verknüpfungstabellen definiert:

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

Dann ist  $(\mathbb{Z}_2, +, \cdot, 0, 1)$  ein Körper, der oft auch mit  $\mathbb{F}_2$  bezeichnet wird. Weiter unten wird dieses Beispiel zu  $\mathbb{Z}_p$  verallgemeinert. ◇

**1.20 Satz** Sei  $(K, +, \cdot, 0, 1)$  ein Körper und seien  $a, b \in K$ . Dann gilt:

- (i)  $ab = 0 \implies a = 0$  oder  $b = 0$  (Nullteilerfreiheit)
- (ii) Die Gleichung  $ax = b$  mit  $a \neq 0$  hat eindeutige Lösung  $x \in K$

**Beweis.**

- (i) Sei  $a \neq 0$  und  $b \neq 0$ , d.h.  $a, b \in K \setminus \{0\}$ . Da  $(K \setminus \{0\}, \cdot)$  eine Gruppe ist, folgt  $a \cdot b \in K \setminus \{0\}$ , d.h.  $ab \neq 0$ . Somit ist die Negation der Aussage gezeigt, und somit auch die Aussage.
- (ii) Falls  $b \neq 0$ , ist die Lösung  $x = a^{-1}b = \frac{b}{a}$  berechnet in der Gruppe  $(K \setminus \{0\}, \cdot)$ . Falls  $b = 0$ , muss nach der Voraussetzung  $a \neq 0$  und (i)  $x = 0$  gelten. □

**1.21 Theorem**  $(\mathbb{Z}_p, +, \cdot, [0], [1])$  Körper  $\iff p \in \mathbb{N} \setminus \{0, 1\}$  Primzahl

**Beweis.** Der Fall  $\mathbb{Z}_2$  wurde schon oben behandelt. Sei also  $p \geq 3$ .

Der Nachweis, dass  $(\mathbb{Z}_p, +, \cdot)$  ein Ring ist, wurde schon geführt. Außerdem ist die Verknüpfung  $\cdot$  in  $(\mathbb{Z}_p \setminus \{[0]\}, \cdot, [1])$  abelsch (d.h. kommutativ) und hat die Eins  $[1]$ . In der Tat gilt  $[n][m] = [nm] = [mn] = [m][n]$  und  $[1][n] = [n]$ . Des Weiteren gilt das Distributivgesetz  $([n] + [m])[k] = [n][k] + [m][k]$ . All dies ist richtig unabhängig von Eigenschaften von  $p$ . Diese sind lediglich für die Existenz des Inversen in  $(\mathbb{Z}_p \setminus \{[0]\}, \cdot, [1])$  relevant.

” $\implies$ ” wird durch einen Widerspruchsbeweis gezeigt. Sei  $p$  keine Primzahl. Dann gibt es  $1 < n, m < p$  so dass  $p = nm$ . Dies impliziert  $[n][m] = [p] = [0]$ , d.h.  $\mathbb{Z}_p$  hat Nullteiler und ist somit nach Satz 1.20(i) kein Körper.

” $\impliedby$ ” Sei  $p$  eine Primzahl. Es ist zu zeigen, dass für  $1 < q < p$  ein Inverses  $[q]^{-1}$  existiert. Setze  $M = \{nq + mp \mid n, m \in \mathbb{Z}, nq + mp > 0\} \subset \mathbb{N}$ . Sei  $d = \min\{M\}$  das kleinste Element der Menge  $M$  bez. der natürlichen Ordnung auf  $\mathbb{Z}$  (die Existenz dieses Minimums folgt, strikt genommen, aus dem Wohlordnungsprinzip, siehe Analysis). Es gilt offensichtlich immer  $d \geq 1$  und  $d \leq p$ , aber wir zeigen nun, dass tatsächlich  $d = 1$  gilt.

*Begründung, dass  $d = 1$ :* Wir beginnen mit der Gegenannahme  $d \geq 2$ . Zerlege  $p = sd + r$  mit Rest  $0 \leq r < d$ . Dann gilt  $r = p - sd = p - s(nq + mp) = (-sn)q + (1 - sm)p \in M$ . Dies impliziert  $r = 0$  nach der Minimalität von  $d$ . Aber  $r = 0$  ist im Widerspruch dazu, dass  $p$  eine Primzahl ist (es wäre  $p = sd$ ). Somit ist  $d = 1$  nachgewiesen.

Zusammenfassend existieren also  $n, m$  mit  $nq + mp = 1$ , d.h.  $[n][q] + [m]p = [1]$ . Da aber  $[m]p = [0]$  folgt  $[q]^{-1} = [n]$ . Somit ist die Existenz von Inversen nachgewiesen. □



**1.22 Bemerkungen** Die Klassifikation endlicher Körper zeigt, dass die Ordnung eines beliebigen endlichen Körpers (d.h. die Anzahl seiner Elemente) immer gleich  $p^n$  ist, wobei  $p$  eine Primzahl ist und  $n \in \mathbb{N} \setminus \{0\}$ . Zudem gilt in  $\mathbb{Z}^p$ , dass  $p[1] = [0]$ . Dies besagt, dass  $\mathbb{Z}_p$  ein Körper von Charakteristik  $p$  ist. Hierbei ist die Charakteristik eines Körpers die kleinste Zahl  $n \in \mathbb{N}$ , so dass  $n \times 1 = 1 + 1 + \dots + 1 = 0$  (hierbei liegen  $n$  Summanden vor). Falls es kein solches  $n$  gibt, spricht man von einem Körper mit Charakteristik 0.  $\diamond$

Die komplexen Zahlen sind das nächste Beispiel eines Körpers.

**1.23 Definition** Als Menge sind die komplexen Zahlen  $\mathbb{C} = \mathbb{R} \times \mathbb{R}$  gegeben durch Paare  $z = (x, y)$  von reellen Zahlen. Auf  $\mathbb{C}$  werden zwei binäre Operationen definiert  $+, \cdot : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$  durch

$$(x, y) + (x', y') = (x + x', y + y'), \quad (x, y) \cdot (x', y') = (xx' - yy', xy' + yx').$$

Dann heißen  $x = \Re(z) \in \mathbb{R}$  und  $y = \Im(z) \in \mathbb{R}$  der Realteil und Imaginärteil von  $z = (x, y) \in \mathbb{C}$ . Anstelle von  $(x, y) \in \mathbb{C}$  wird meist  $z = x + iy$  geschrieben, wobei  $i$  die sogenannte imaginäre Einheit ist und zunächst lediglich ein Symbol ist, was den Imaginärteil spezifiziert. Für diese Korrespondenz schreiben wir  $x + iy \cong (x, y)$ .

**1.24 Bemerkungen** Seien  $z = x + iy$  und  $z' = x' + iy'$ . Dann werden obige Verknüpfungen zu

$$z + z' = (x + x') + i(y + y'), \quad z \cdot z' = (x + iy)(x' + iy') = (xx' - yy') + i(xy' + yx').$$

Insbesondere gilt also für die imaginäre Einheit  $i \cong (0, 1)$

$$i^2 = i \cdot i = (0 - 1) + i(0 + 0) = -1,$$

Anders ausgedrückt,  $i$  löst die Gleichung  $z^2 = -1$ , die keine reelle Lösung hat. Dass es in der Menge der komplexen Zahlen eine solche Lösung gibt, ist ein großer Vorteil, der ihre Einführung rechtfertigt. Etwas mehr dazu weiter unten (Fundamentalsatz der Algebra). Die komplexe Null und die komplexe Eins sind  $0_{\mathbb{C}} \cong (0, 0)$  und  $1_{\mathbb{C}} \cong (1, 0)$ . Sie erfüllen nach obigen Regeln

$$0_{\mathbb{C}} + z = z, \quad 1_{\mathbb{C}} \cdot z = z,$$

für alle  $z \in \mathbb{C}$ . Dies sind schon Vorbereitungen für den folgenden Satz.  $\diamond$

**1.25 Satz**  $(\mathbb{C}, +, \cdot, 0_{\mathbb{C}}, 1_{\mathbb{C}})$  ist ein Körper. Die reellen Zahlen bilden einen Unterkörper von  $\mathbb{C}$ .

**Beweis.** Zunächst ist  $(\mathbb{C}, +, 0_{\mathbb{C}})$  eine abelsche Gruppe mit inversem  $z^{-1} = -z$ . Als Nächstes wird überprüft, dass  $(\mathbb{C}, \cdot, 1_{\mathbb{C}})$  auch eine abelsche Gruppe ist. Die Assoziativität muss nachgerechnet werden:

$$\begin{aligned} (x + iy) \cdot \{(x' + iy') \cdot (x'' + iy'')\} &= (x + iy) \cdot (x'x'' - y'y'' + i(x'y'' + y'x'')) \\ &= \dots = \{(x + iy)(x' + iy')\} \cdot (x'' + iy''). \end{aligned}$$

Außerdem ist das Inverse gegeben durch

$$(x + iy)^{-1} = \frac{x}{x^2 + y^2} - i \frac{y}{x^2 + y^2},$$

wie man wieder nachrechnen kann. Zu guter Letzt muss auch das Distributivgesetz verifiziert werden (Übung). Die Unterkorpereigenschaft von  $\mathbb{R}$  ist offensichtlich.  $\square$

**1.26 Bemerkung** Von nun an werden wie allgemein üblich die Indizes von  $0_{\mathbb{C}} = 0$  und  $1_{\mathbb{C}} = 1$  weggelassen, ebenso wie der Punkt  $\cdot$  für die Multiplikation.  $\diamond$

**1.27 Definition** Zu jeder komplexen Zahl  $z = x + iy \in \mathbb{C}$  sind ihr komplex Konjugiertes  $\bar{z} \in \mathbb{C}$ , Betrag  $|z| \in \mathbb{R}_{\geq}$  und Argument  $\arg(z) \in [0, 2\pi)$  definiert durch

$$\bar{z} = x - iy, \quad |z| = \sqrt{x^2 + y^2}, \quad \arg(z) = \arctan\left(\frac{y}{x}\right).$$

Außerdem setzen wird für einen Winkel  $\varphi \in [0, 2\pi)$ , bzw. sogar für  $\varphi \in \mathbb{R}$ ,

$$e^{i\varphi} = \cos(\varphi) + i \sin(\varphi), \quad e^z = e^{\operatorname{Re}(z)} e^{i\operatorname{Im}(z)}. \quad (1.1)$$

Hierbei sind die trigonometrischen Funktionen  $\cos$ ,  $\sin$  und  $\arctan$  als bekannt vorausgesetzt, ebenso wie die reelle Exponentialfunktion als die eindeutige bestimmte Funktion  $e : \mathbb{R} \rightarrow \mathbb{R}_{>}$  mit  $e^{x+x'} = e^x e^{x'}$  und  $e^0 = 1$  (Schule bzw. Analysis).

**1.28 Bemerkungen** 1. In der Analysisvorlesung wird die komplexe Exponentialfunktion  $z \in \mathbb{C} \mapsto e^z$  durch eine konvergente Reihe definiert und aus dieser werden dann die hier als Definition verwandten Gleichungen (1.1) hergeleitet. Der hier verwandte *ad hoc* Zugang ist aber auch in sich schlüssig und erlaubt uns schon jetzt mit den komplexen Zahlen zu arbeiten.

2. Es seien einige Eigenschaften von der komplexen Exponentialfunktion aufgelistet. Zunächst besagt der Satz von Pythagoras

$$|e^{i\varphi}| = \sqrt{\cos^2(\varphi) + \sin^2(\varphi)} = 1.$$

Außerdem gilt für das komplexe Konjugat:

$$\overline{e^{i\varphi}} = e^{-i\varphi}.$$

Durch Einsetzen spezieller Werte in die trigonometrischen Funktionen erhält man folgende Identitäten:

$$e^{i0} = 1 = e^{i2\pi}, \quad e^{i\pi} = -1, \quad e^{i\frac{\pi}{2}} = i, \quad e^{i\frac{3\pi}{2}} = -i.$$

Desweiteren folgt mit den Additionstheoremen der Trigonometrie, dass für  $\varphi, \psi \in \mathbb{R}$

$$\begin{aligned} e^{i\varphi} e^{i\psi} &= (\cos(\varphi) + i \sin(\varphi))(\cos(\psi) + i \sin(\psi)) \\ &= (\cos(\varphi) \cos(\psi) - \sin(\varphi) \sin(\psi)) + i (\cos(\psi) \sin(\varphi) + \cos(\varphi) \sin(\psi)) \\ &= \cos(\varphi + \psi) + i \sin(\varphi + \psi) \\ &= e^{i(\varphi + \psi)}. \end{aligned}$$

Umgekehrt können aus der Identität  $e^{i\varphi} e^{i\psi} = e^{i(\varphi + \psi)}$  die Additionstheoreme hergeleitet werden (mnemotechnisch ist wohl einfacher sich diese Identität zu merken). Des Weiteren kann nun, soweit  $e^{x+x'} = e^x e^{x'}$  für die reelle Exponentialfunktion bekannt ist, verifiziert werden, dass

$$e^{z+z'} = e^z e^{z'}, \quad z, z' \in \mathbb{C}.$$

Wie oben schon angedeutet, folgt diese Gleichung auch aus der Reihendarstellung der komplexen Exponentialfunktion (siehe Analysis).

3. Die Polardarstellung einer komplexen Zahl  $z \in \mathbb{C}$  ist

$$z = r e^{i\varphi} = r \cos(\varphi) + i r \sin(\varphi), \quad r = |z|, \quad \varphi = \arg(z).$$

Dies folgt direkt aus den trigonometrischen Identitäten

$$\cos(\arctan(\frac{x}{y})) = \frac{x}{\sqrt{x^2 + y^2}}, \quad \sin(\arctan(\frac{x}{y})) = \frac{y}{\sqrt{x^2 + y^2}}.$$

In der Polardarstellung von  $z = r e^{i\varphi}$  und  $z' = r' e^{i\varphi'}$  ist die Multiplikation komplexer Zahlen

$$z z' = r r' e^{i(\varphi + \varphi')}.$$

Insbesondere gilt also  $|z z'| = |z| |z'|$  und  $\arg(z z') = \arg(z) + \arg(z')$ .

4. Die komplexe Konjugation erfüllt, für  $z, w \in \mathbb{C}$ ,

$$\overline{z + w} = \bar{z} + \bar{w}, \quad \overline{z w} = \bar{z} \bar{w}, \quad \overline{\bar{z}} = z, \quad z \bar{z} = |z|^2,$$

sowie

$$\frac{1}{2}(z + \bar{z}) = \Re(z), \quad \frac{1}{2i}(z - \bar{z}) = \Im(z).$$

◇

**1.29 Satz (Eigenschaften des Betrages)** Seien  $z, w \in \mathbb{C}$ .

- (i)  $|z| \geq 0$  und ( $|z| = 0 \iff z = 0$ )
- (ii)  $|z + w| \leq |z| + |w|$  (Dreiecksungleichung)
- (iii)  $|z w| = |z| |w|$
- (iv)  $|\bar{z}| = |z|$
- (v)  $\max\{|\Re(z)|, |\Im(z)|\} \leq |z| \leq |\Re(z)| + |\Im(z)|$

**Beweis.** (i) und (iv) sind offensichtlich, (iii) schon oben verifiziert. Zu (ii):

$$\begin{aligned} |z + w| &= \left| |z| e^{i\arg(z)} + |w| e^{i\arg(w)} \right| \\ &= \left| |z| e^{i\varphi} + |w| e^{-i\varphi} \right| \quad \text{mit } \varphi = \frac{1}{2}(\arg(z) - \arg(w)) \\ &= \sqrt{(|z| + |w|)^2 \cos^2(\varphi) + (|z| - |w|)^2 \sin^2(\varphi)} \\ &\leq (|z| + |w|) \sqrt{\cos^2(\varphi) + \sin^2(\varphi)} = |z| + |w|. \end{aligned}$$

Später wird für diese Ungleichung ein wesentlich besserer, d.h. strukturbetonender, Beweis angegeben. Zu (v) für  $z = x + iy$ :

$$\max\{|x|, |y|\} \leq \sqrt{x^2 + y^2} \leq |x| + |y|,$$

Letzteres, da  $x^2 + y^2 \leq x^2 + y^2 + 2|x||y|$ . □

Wichtigster Grund für die Einführung der komplexen Zahlen ist folgender Sachverhalt, der in der Analysisvorlesung oder einer Vorlesung über Funktionentheorie bewiesen wird.

**1.30 Satz (Fundamentalsatz der Algebra)** Sei  $p$  ein Polynom  $N$ -ten Grades,  $N \in \mathbb{N}$ , mit komplexen Koeffizienten:

$$p(z) = \sum_{n=0}^N p_n z^n, \quad p_n \in \mathbb{C}, \quad p_N \neq 0.$$

Dann hat  $p$  genau  $N$  Nullstellen  $z_1, \dots, z_N \in \mathbb{C}$ , d.h.

$$p(z) = p_N(z - z_1)(z - z_2) \cdot \dots \cdot (z - z_N).$$

**1.31 Bemerkung** Der Satz ist lediglich eine sogenannte Existenzaussage für die  $N$  Nullstellen des Polynomes  $p$  (auch Nullstellen genannt). Der Beweis konstruiert die Nullstellen nicht und liefert keine explizite Lösungsformel. So eine Lösungsformel für Polynome zweiten Grades

$$p(z) = z^2 + az + b,$$

ist aus der Schule bekannt, kann aber auch direkt verifiziert werden:

$$z_1 = -\frac{a}{2} + \sqrt{\frac{a^2}{4} - b}, \quad z_2 = -\frac{a}{2} - \sqrt{\frac{a^2}{4} - b}.$$

Für Polynome dritten und vierten Grades gibt es auch Lösungsformeln, die aber schon wesentlich schwieriger aussehen. Für Polynome höheren Grades kann bewiesen werden, dass es im Allgemeinen keine Lösungsformel gibt.  $\diamond$

## 2 Die Struktur des Vektorraums

Von nun an bezeichnet  $\mathbb{K}$  einen Körper. In allen Beispielen wird entweder  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$  sein.

**2.1 Definition** Eine abelsche Gruppe  $(V, +, 0)$  heißt Vektorraum über  $\mathbb{K}$ , falls sie mit einer sogenannten Skalarmultiplikation  $\cdot : \mathbb{K} \times V \rightarrow V$  versehen ist, welche für alle Skalare  $\lambda, \mu \in \mathbb{K}$  und Vektoren  $v, w \in V$  folgende Axiome erfüllt:

- (i) Pseudoassoziativität:  $\lambda \cdot (\mu \cdot v) = (\lambda\mu) \cdot v$
- (ii) Skalare Distributivität:  $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$
- (iii) Vektordistributivität:  $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$
- (iv) Normierung:  $1_{\mathbb{K}} \cdot v = v$

Dann heißt  $0 \in V$  auch der Nullvektor.

Falls  $\mathbb{K} = \mathbb{R}$ , heißt  $V$  ein reeller Vektorraum, und falls  $\mathbb{K} = \mathbb{C}$  ein komplexer Vektorraum.

Notation: Der Punkt  $\cdot$  wird meist weggelassen, denn  $\lambda v$  für eine Zahl  $\lambda \in \mathbb{K}$  und einen Vektor  $v \in V$  kann nichts anderes als die Skalarmultiplikation bezeichnen.

**2.2 Beispiele** 1. Sei  $V = \{0\}$  die triviale Gruppe. Durch  $\lambda 0 = 0$  wird dies zu einem Vektorraum über  $\mathbb{K}$ , den man null-dimensional nennt.

2. Sei  $V = \mathbb{K}$ . Darauf wird die Multiplikation in  $\mathbb{K}$  als Skalarmultiplikation mit Elementen aus  $\mathbb{K}$  aufgefaßt. Damit wird  $V = \mathbb{K}$  zu einem Vektorraum über  $\mathbb{K}$ , den man als den eindimensionalen Vektorraum über  $\mathbb{K}$  bezeichnet.

3. Nun sei  $N \in \mathbb{N}$  und betrachte  $\mathbb{K}^N = \mathbb{K} \times \mathbb{K} \times \dots \times \mathbb{K}$ , wobei in letzterem mengentheoretischen Produkt  $N$  Faktoren vorliegen. Ein Vektor  $v = (v_1, \dots, v_N) \in \mathbb{K}^N$  besitzt also  $N$  sogenannte Komponenten  $v_n \in \mathbb{K}$  für  $n = 1, \dots, N$ . Meist werden Vektoren im  $\mathbb{K}^N$  nicht wie eben als Zeile geschrieben, sondern eher als Spalte:

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix},$$

aber dies ist lediglich eine Wahl, die allerdings später für die Matrix-Vektor-Verknüpfung sehr vorteilhaft ist. Die Vektoraddition und Skalarmultiplikation mit  $\lambda \in \mathbb{K}$  sind nun wie folgt gegeben:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} + \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_N + w_N \end{pmatrix}, \quad \lambda \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} = \begin{pmatrix} \lambda v_1 \\ \vdots \\ \lambda v_N \end{pmatrix}.$$

Oft werden diese beiden Definitionen auch zu einer zusammengefaßt:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} + \lambda \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix} = \begin{pmatrix} v_1 + \lambda w_1 \\ \vdots \\ v_N + \lambda w_N \end{pmatrix}.$$

Die erste Gleichung entspricht dann dem Fall  $\lambda = 1$ , die zweite  $w = 0$  (Nullvektor, d.h.  $w_1 = \dots = w_N = 0_{\mathbb{K}}$ ). Tatsächlich sind mit diesen Definitionen alle Axiome eines Vektorraumes erfüllt. Man bezeichnet  $\mathbb{K}^N$  auch als den  $N$ -dimensionalen Vektorraum über  $\mathbb{K}$ . Eine abstraktere Version des Dimensionsbegriffes folgt.

Aus der Schule bekannt sein könnten die Spezialfälle  $\mathbb{K} = \mathbb{R}$  und  $N = 2$  oder  $N = 3$ . Der zweidimensionale reelle Vektorraum  $\mathbb{R}^2$  besteht aus allen Vektoren  $v = \begin{pmatrix} x \\ y \end{pmatrix}$  mit reellen Koeffizienten  $x, y \in \mathbb{R}$  (beachte, dass  $\mathbb{R}^2$  auch mit der Menge der komplexen Zahlen identifiziert werden kann). Seine Elemente (Vektoren) können als Pfeile dargestellt werden, die vom Ursprung zum Punkt  $(x, y)$  hinzeigen. Die Addition von Vektoren ist dann das *Aneinanderkleben von Vektoren*, die Skalarmultiplikation mit Skalaren aus  $\mathbb{R}$  die Streckung oder Stauchung eines Vektors. Analoges gilt im  $\mathbb{R}^3$ , nur dass die Vektoren nun 3 Komponenten haben.

4. Im vorigen Beispiel tauchte  $\mathbb{C}^N$  als Vektorraum über  $\mathbb{C}$  auf, was man auch kurz als komplexen Vektorraum bezeichnet. Es ist jedoch möglich, die Menge  $\mathbb{C}^N$  auch als reellen Vektorraum aufzufassen, indem in der Skalarmultiplikation lediglich die Teilmenge  $\mathbb{R} \subset \mathbb{C}$  als Skalare verwandt wird. Es wird dann von dem reellen Vektorraum  $\mathbb{C}^N$  gesprochen.
5. Sei  $\mathbb{K}[x]$  die Menge der Polynome in  $x$  mit Koeffizienten in  $\mathbb{K}$ , auf der genau wie in Beispiel 1.15 eine Addition definiert ist (durch komponentenweises Addieren). Nun wird zudem eine Skalarmultiplikation definiert durch

$$(\lambda p)(x) = \sum_{n=0}^N (\lambda p_n) x^n, \quad \lambda \in \mathbb{K}, \quad p(x) = \sum_{n=0}^N p_n x^n.$$

Hierbei sind  $p_n \in \mathbb{K}$ ,  $n = 1, \dots, N$ , die Koeffizienten von  $p$ . Wiederum überprüft man direkt, dass  $\mathbb{K}[x]$  zu einem Vektorraum wird. Seine Dimension ist unendlich. Wenn man hingegen die Menge  $\mathbb{K}_N[x]$  der Polynome vom Grad kleiner oder gleich  $N$  betrachtet, so bildet  $\mathbb{K}_N[x]$  versehen mit den gleichen Operationen auch einen Vektorraum, dessen Dimension  $N + 1$  ist (denn es liegen  $N + 1$  freie Parameter  $p_0, p_1, \dots, p_N$  vor).

6. Sei  $X$  eine beliebig Menge und  $F(X, \mathbb{K})$  die Menge der Abbildungen von  $X$  nach  $\mathbb{K}$ . Für  $f, g \in F(X, \mathbb{K})$  und  $\lambda \in \mathbb{K}$  setzen wir

$$(f + \lambda g)(x) = f(x) + \lambda g(x), \quad x \in X.$$

Beachte hierbei, dass auf der rechten Seite die Operationen in  $\mathbb{K}$  verwandt werden, um auf der linken Seite neue Strukturen auf  $F(X, \mathbb{K})$  zu definieren. Wiederum wird  $F(X, \mathbb{K})$  somit zu einem Vektorraum über  $\mathbb{K}$ .  $\diamond$

**2.3 Definition** Eine Teilmenge  $U \subset V$  eines Vektorraumes  $V$  über  $\mathbb{K}$  heißt *Unterraum*, genau dann, wenn  $U$  Untergruppe von  $(V, +, 0)$  ist und  $\lambda u \in U$  für alle  $\lambda \in \mathbb{K}$  und  $u \in U$ , d.h. genau dann, wenn  $U$  mit den von  $V$  vererbten Operationen wieder ein Vektorraum ist.

**2.4 Bemerkung** Um zu verifizieren, dass  $U$  ein Unterraum von  $V$  ist, reicht es zu zeigen, dass  $u + \lambda w \in U$  für alle  $u, w \in U$  und  $\lambda \in \mathbb{K}$  und  $U$  nicht-leer ist (also insbesondere den Nullvektor enthält, denn mit  $u \in U$  ist auch  $u + (-1)u = 0 \in U$ ).  $\diamond$

**2.5 Beispiele** 1. Im Vektorraum  $\mathbb{R}^2$  sind die Linien durch den Ursprung von Steigung  $\alpha \in \mathbb{R}$

$$U = \left\{ v = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid y = \alpha x \right\},$$

Unterräume. Analog sind Linien und Ebenen im  $\mathbb{R}^3$ , die den Ursprung enthalten, auch Unterräume des  $\mathbb{R}^3$ . Allgemeiner bilden sogenannte Hyperebenen im  $\mathbb{R}^N$ , die den Ursprung enthalten, genau die Unterräume des  $\mathbb{R}^N$ . Mehr dazu später.

2. Die Polynome  $\mathbb{K}_N[x]$  vom Grad kleiner oder gleich  $N \in \mathbb{N}$  bilden einen Unterraum von  $\mathbb{K}[x]$ .

3. Die Polynome  $\mathbb{K}[x]$  in  $x \in \mathbb{R}$  bilden einen Unterraum des Vektorraumes  $F(\mathbb{R}, \mathbb{K})$  der Funktionen von  $\mathbb{R}$  nach  $\mathbb{K}$ .  $\diamond$

**2.6 Satz** Der Durchschnitt von beliebig vielen Unterräumen ist wieder ein Unterraum.

**Beweis.** Sei  $\mathcal{U} = (U)_{U \in \mathcal{U}}$  eine Familie von Unterräumen eines Vektorraumes  $V$ . Dann enthält der Durchschnitt  $D = \bigcap_{U \in \mathcal{U}} U$  den Nullvektor, weil er ja in jedem Unterraum enthalten ist. Außerdem seien  $v, w \in D$  und  $\lambda \in \mathbb{K}$ . Dann sind  $v, w \in U$  für alle  $U \in \mathcal{U}$ . Da jedes  $U$  ein Unterraum ist, folgt, dass  $v + \lambda w \in U$  für jedes  $U \in \mathcal{U}$ . Somit ist  $v + \lambda w \in D$ , d.h.  $D$  ist ein Vektorraum.  $\square$

**2.7 Bemerkung** Die Vereinigung von zwei Unterräumen ist nur dann ein Unterraum, wenn einer der Unterräume alle anderen enthält. Z.B. betrachten wir  $\mathbb{C}$  als reellen Vektorraum. Dann sind die reelle Achse  $\mathbb{R} \subset \mathbb{C}$  und imaginäre Achse  $i\mathbb{R} = \{iy \mid y \in \mathbb{R}\} \subset \mathbb{C}$  zwei Unterräume. Aber die Vereinigung  $\mathbb{R} \cup i\mathbb{R}$  ist das Achsenkreuz und somit kein Vektorraum, denn z.B.  $1, i \in \mathbb{R} \cup i\mathbb{R}$ , aber  $1 + i \notin \mathbb{R} \cup i\mathbb{R}$ . Es ist jedoch möglich, aus der Vereinigung von Unterräumen auf natürliche Art und Weise einen Unterraum zu konstruieren. Dies wird in der nächsten Definition getan.  $\diamond$

**2.8 Definition** Gegeben seien  $N$  Vektoren  $v_1, \dots, v_N$  in einem Vektorraum  $V$  über  $\mathbb{K}$ . Eine *Linearkombination* dieser Vektoren ist dann

$$\lambda_1 v_1 + \dots + \lambda_N v_N = \sum_{n=1}^N \lambda_n v_n,$$

wobei  $\lambda_1, \dots, \lambda_N \in \mathbb{K}$ . Der Spann der  $N$  Vektoren ist definiert als die Menge aller Linearkombinationen:

$$\text{span}_{\mathbb{K}}(\{v_1, \dots, v_N\}) = \left\{ \sum_{n=1}^N \lambda_n v_n \mid \lambda_1, \dots, \lambda_N \in \mathbb{K} \right\} .$$

Etwas allgemeiner wird auch für eine beliebige, nicht notwendigerweise endliche Menge  $M \subset V$  der Spann definiert:

$$\text{span}_{\mathbb{K}}(M) = \{ \text{Linearkombinationen von endlich vielen Vektoren aus } M \} .$$

Falls  $\text{span}_{\mathbb{K}}(M) = V$  heißt  $M$  ein Erzeugendensystem von  $V$ .

Falls  $U, U'$  Unterräume von  $V$  sind, wird  $\text{span}_{\mathbb{K}}(U \cup U')$  auch mit  $U + U'$  bezeichnet.

**2.9 Satz** Für jede Teilmenge  $M \subset V$  eines Vektorraumes ist  $\text{span}_{\mathbb{K}}(M)$  ein Unterraum. Er ist auch gegeben durch

$$\text{span}_{\mathbb{K}}(M) = \bigcap_{U, M \subset U} U , \quad (2.1)$$

wobei der Durchschnitt gebildet wird über alle Unterräume  $U$  von  $V$ , die die Menge  $M$  enthalten.

**Beweis.** Es ist zu zeigen, dass wenn  $v, w \in \text{span}_{\mathbb{K}}(M)$  und  $\lambda \in \mathbb{K}$ , dann ist auch  $v + \lambda w \in \text{span}_{\mathbb{K}}(M)$ . Seien also

$$v = \sum_{n=1}^N \eta_n v_n , \quad w = \sum_{k=1}^K \mu_k w_k ,$$

wobei  $N, K \in \mathbb{N}$  und  $\eta_n, \mu_k \in \mathbb{K}$  und  $v_n, w_k \in M$ . Dann ist

$$v + \lambda w = \sum_{n=1}^N \eta_n v_n + \sum_{k=1}^K (\lambda \mu_k) w_k ,$$

was also wieder eine endliche Linearkombination von Vektoren aus  $M$  ist, d.h.  $v + \lambda w \in \text{span}_{\mathbb{K}}(M)$ .

Es verbleibt noch die Gleichheit (2.1) zu zeigen. Sei  $D = \bigcap_{U, M \subset U} U$  der Durchschnitt auf der rechten Seite. Dann ist offensichtlich  $M \subset D$  und, da  $D$  nach Satz 2.6 ein Unterraum ist, folgt auch, dass  $\text{span}_{\mathbb{K}}(M) \subset D$ . Andererseits ist nach dem eben Bewiesenen auch  $\text{span}_{\mathbb{K}}(M)$  ein Unterraum, der offensichtlich  $M$  enthält. Also ist er einer der Unterräume die im Durchschnitt  $D$  auftreten und es folgt auch die umgekehrte Inklusion  $D \subset \text{span}_{\mathbb{K}}(M)$ .  $\square$

**2.10 Beispiele** 1. Im reellen Vektorraum  $\mathbb{R}^3$  seien zwei Vektoren  $v, w \in \mathbb{R}^3$  gegeben. Dann ist

$$\text{span}_{\mathbb{R}}(\{v, w\}) = \{ \lambda v + \mu w \mid \lambda, \mu \in \mathbb{R} \} ,$$

der Unterraum, der durch die Ebene gegeben wird, die die Vektoren  $v$  und  $w$  enthält (und den Ursprung). Falls  $v = \eta w$  für ein  $\eta \in \mathbb{R}$ , so liegt nicht eine Ebene, sondern eine Gerade durch den Ursprung vor.

2. Im reellen Vektorraum  $\mathbb{R}[x]$  der Polynome in  $x$  ist der Spann der Vektoren  $p(x) = 3x$  und  $q(x) = 2x^4 + 1$  gegeben durch

$$\text{span}_{\mathbb{R}}(\{p, q\}) = \{ 2\lambda x^4 + 3\mu x + \lambda \mid \lambda, \mu \in \mathbb{R} \} .$$

3. Nun greifen wir noch einmal das Beispiel des reellen Vektorraumes  $\mathbb{C}$  aus Bemerkung 2.7 auf, in dem zwei Unterräume  $\mathbb{R}$  und  $i\mathbb{R}$  gegeben sind. Dann ist der Spann von  $M = \mathbb{R} \cup i\mathbb{R}$  gleich ganz  $\mathbb{C}$ . Dies besagt, dass  $M$  ein Erzeugendensystem ist. Es gibt jedoch auch viel kleinere Erzeugendensysteme, z.B.  $M' = \{1, i\}$  enthält nur 2 Punkte, aber tatsächlich gilt  $\text{span}_{\mathbb{R}}(M') = \mathbb{C}$ . Kleinstmögliche (minimale) Erzeugendensysteme heißen auch Basen. Dieser Begriff wird weiter unten definiert.  $\diamond$

**2.11 Definition** Sei  $V$  ein Vektorraum über  $\mathbb{K}$ .

- (i) Vektoren  $v_1, \dots, v_N \in V$  heißen *linear unabhängig über  $\mathbb{K}$* , genau dann, wenn

$$\sum_{n=1}^N \lambda_n v_n = 0 \quad \text{mit } \lambda_1, \dots, \lambda_N \in \mathbb{K} \quad \implies \quad \lambda_1 = \dots = \lambda_N = 0,$$

d.h. genau dann, wenn der Nullvektor sich nur durch eine triviale Linearkombination der Vektoren  $v_1, \dots, v_N$  darstellen lässt.

- (ii) Vektoren  $v_1, \dots, v_N \in V$  heißen *linear abhängig*, genau dann, wenn sie nicht linear unabhängig sind.  
 (iii) Eine Teilmenge  $M \subset V$  heißt *linear unabhängig*, genau dann, wenn alle endlich viele Vektoren  $v_1, \dots, v_N \in M$  linear unabhängig sind.

**2.12 Beispiele** 1. Beginnen wir mit einem ganz konkreten Beispiel dreier Vektoren im reellen Vektorraum  $\mathbb{R}^3$ :

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Dann sind  $v_1, v_2, v_3$  linear abhängig weil  $v_1 - v_2 + v_3 = 0$ . Andererseits sind  $v_1, v_2$  linear unabhängig, denn

$$\lambda_1 v_1 + \lambda_2 v_2 = \begin{pmatrix} \lambda_1 + 2\lambda_2 \\ 2\lambda_1 + 3\lambda_2 \\ 3\lambda_1 + 4\lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

impliziert, unter Verwendung der drei Gleichungen für die drei Komponenten,  $\lambda_1 = \lambda_2 = 0$ . Ein systematischer Zugang zur Überprüfung der linearen Abhängigkeit bzw. Unabhängigkeit ist der Gauß-Algorithmus, der bald vorgestellt wird. Analog sind auch  $v_1, v_3$  linear unabhängig und  $v_2, v_3$  linear unabhängig. Geometrisch bedeutet dies, dass die drei Vektoren  $v_1, v_2, v_3$  in einer zweidimensionalen Ebene liegen, die von je zwei der beteiligten Vektoren aufgespannt wird. Diese Ebene bildet einen Unterraum.

2. Im komplexen Vektorraum  $\mathbb{C}^2$  sind

$$v_1 = \begin{pmatrix} 1 \\ 1+i \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2i \\ -2+2i \end{pmatrix},$$

linear abhängig. Tatsächlich gilt  $2iv_1 = v_2$ , d.h.  $2iv_1 - v_2 = 0$ . Wenn hingegen  $\mathbb{C}^2$  als reeller Vektorraum aufgefaßt wird, so sind  $v_1, v_2$  linear unabhängig, denn es gibt keine reellen Zahlen  $\lambda_1, \lambda_2 \in \mathbb{R}$  mit  $\lambda_1 v_1 + \lambda_2 v_2 = 0$  bis auf  $\lambda_1 = \lambda_2 = 0$ .

3. Mengen  $M \subset V$ , die den Nullvektor enthalten, sind immer linear abhängig. Leere Mengen hingegen sind linear unabhängig.



4. Einelementige Mengen sind immer linear unabhängig, es sei denn sie bestehen aus dem Nullvektor.
5. Teilmengen linear unabhängiger Mengen sind linear unabhängig.
6. Obermengen linear abhängiger Mengen sind linear abhängig. ◇

**2.13 Definition** Sei  $V$  ein Vektorraum über  $\mathbb{K}$ . Ein linear unabhängiges Erzeugendensystem  $\mathcal{B} \subset V$  heißt eine Basis. Die Dimension  $\dim_{\mathbb{K}}(V) \in \mathbb{N}_0 \cup \{\infty\}$  des Vektorraumes ist definiert als die Anzahl der Elemente von  $\mathcal{B}$ .

- 2.14 Bemerkungen**
- (i) Wenn eine Menge (wie der  $\mathbb{C}^N$ ) sowohl die Struktur eines komplexen wie reellen Vektorraumes hat, so ist es nötig zu spezifizieren bez. welcher Struktur die Dimension berechnet wird. Es wird deswegen von sogenannter reeller Dimension und komplexer Dimension gesprochen, allgemeiner von einer  $\mathbb{K}$ -Dimension  $\dim_{\mathbb{K}}$ . Der Unterschied wird weiter unten in einem Beispiel erläutert.
  - (ii) Die Dimension ist gleich der Anzahl der freien Parameter (in  $\mathbb{K}$ ), die benötigt werden, um einem Vektor darzustellen.
  - (iii) Damit die Definition der Dimension wirklich einen Sinn macht, muss nachgewiesen werden, dass jeder Vektorraum eine Basis hat und dass jede Basis gleich viele Elemente hat. Dies wird letztendlich gezeigt (Korollar 2.20), aber vielleicht fangen wir erst einmal mit Beispielen an. ◇

**2.15 Beispiele** 1. Wir betrachten im reellen Vektorraum  $\mathbb{R}^N$  die Vektoren

$$b_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots \quad b_N = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Sie sind offensichtlich linear unabhängig und jeder Vektor  $v$  kann als Linearkombination von  $b_1, \dots, b_N$  geschrieben werden:

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} \implies v = \sum_{n=1}^N v_n b_n. \quad (2.2)$$

Dies bedeutet in der Tat, dass  $\text{span}_{\mathbb{R}}(\{b_1, \dots, b_N\}) = \mathbb{R}^N$ , so dass  $b_1, \dots, b_N$  eine Basis des  $\mathbb{R}^N$  bilden. Die reelle Dimension von  $\mathbb{R}^N$  ist also  $N$ .

2. Es sei schon hier betont, dass es zwar eine Standardbasis des  $\mathbb{R}^N$  gibt, aber noch viele andere. Z.B. bilden die Vektoren

$$b'_n = \sum_{k=1}^n b_k$$

auch eine Basis des  $\mathbb{R}^N$ . Dies zu überprüfen ist eine Übung, die mit adhoc Methoden durchgeführt werden kann. Ein allgemeines Vorgehen zum Nachweis der Basiseigenschaften wird im Zusammenhang mit dem Gauß-Verfahren vorgestellt.

3. Die gleichen Vektoren  $b_1, \dots, b_N$  bilden auch eine Basis des komplexen Vektorraumes  $\mathbb{C}^N$ . In der Tat, die Zerlegung (2.2) ist immer noch richtig, nur dass die Koeffizienten  $v_n$  nun komplexe Zahlen sind. Somit ist die komplexe Dimension von  $\mathbb{C}^N$  gleich  $N$ , d.h.  $\dim_{\mathbb{C}}(\mathbb{C}^N) = N$ .
4. Andererseits bilden  $b_1, \dots, b_N$  keine Basis des reellen Vektorraumes  $\mathbb{C}^N$ , denn (2.2) ist zwar richtig, aber die rechte Seite stellt keine reelle Linearkombination dar, in der ja nur reelle Koeffizienten verwendet werden dürfen. Es ist jedoch möglich die Menge  $b_1, \dots, b_N$  zu einer Basis zu ergänzen durch  $b_{N+n} = ib_n$  für  $n = 1, \dots, N$ . Dann gilt

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} = \sum_{n=1}^N (\Re(v_n) b_n + \Im(v_n) b_{n+N}) .$$

Somit hat  $\mathbb{C}^N$  reelle Dimension  $2N$ , d.h.  $\dim_{\mathbb{R}}(\mathbb{C}^N) = 2N$ .

5. Nun betrachten wir den  $\mathbb{K}$ -Vektorraum  $\mathbb{K}[x]$  der Polynome in  $x$  mit Koeffizienten in  $\mathbb{K}$ . Dann bilden die sogenannten Monome

$$b_n(x) = x^n ,$$

ein Basis  $\mathcal{B} = \{b_n \mid n \geq 0\}$ , die auch mit  $\mathcal{B} = (b_n)_{n \geq 0}$  bezeichnet wird. Da diese Basis unendlich viele Elemente hat, ist die  $\mathbb{K}$ -Dimension von  $\mathbb{K}[x]$  gleich unendlich  $\infty$ . Für den Vektorraum  $\mathbb{K}_N[x]$  der Polynome vom Grad kleiner oder gleich  $N$  ist die  $\mathbb{K}$ -Dimension  $N + 1$  (geben Sie eine entsprechende Basis an).  $\diamond$

Nach diesen erläuternden Beispielen werden nun allgemeine Sachverhalte nachgewiesen.

### 2.16 Satz Jeder Vektorraum besitzt eine Basis.

**Beweis.** Es ist offensichtlich, dass jeder Vektorraum ein Erzeugendensystem hat (insbesondere, einfach der Vektorraum selber), aber nun soll ja ein Erzeugendensystem aus linear unabhängigen Vektoren konstruiert werden. Hierzu kann mit einem ersten Vektor  $b_1 \neq 0$  begonnen werden, es sei denn der Vektorraum  $V$  besteht nur aus dem Nullvektor. Falls  $\text{span}_{\mathbb{K}}(\{b_1\}) = V$  gilt, so ist schon eine Basis gefunden. Falls nicht, d.h.  $\text{span}(\{b_1\}) \neq V$ , gibt es ja einen Vektor  $b_2 \in V \setminus \text{span}_{\mathbb{K}}(\{b_1\})$ , der also von nicht als Linearkombination von  $b_1$  dargestellt werden kann und somit sind  $b_1, b_2$  linear unabhängig. Nun iteriert man das Verfahren. Im  $N$ ten Schritt liegen  $b_1, \dots, b_N$  vor. Entweder  $\text{span}_{\mathbb{K}}(\{b_1, \dots, b_N\}) = V$  und die Konstruktion ist beendet, oder es gibt ein  $b_{N+1} \notin \text{span}_{\mathbb{K}}(\{b_1, \dots, b_N\})$ . In einem endlich-dimensionalen Vektorraum bricht die Konstruktion ab, aber in einem unendlich dimensionalen Vektorraum muß unendlich oft iteriert werden. Mathematisch sauber wird an dieser Stelle mit dem sogenannten Zorn'schen Lemma argumentiert. Alle daran Interessierten sind eingeladen, dies in die Büchern der Literaturliste nachzulesen.  $\square$

### 2.17 Satz (Basisergänzungssatz) Jede linear unabhängige Teilmenge $M \subset V$ eines Vektorraumes kann zu einer Basis ergänzt werden.

**Beweis.** Dies folgt aus dem Beweis des letzten Satzes, wenn gleich mit  $M$  begonnen wird.  $\square$

### 2.18 Satz Sei $V$ ein Vektorraum über $\mathbb{K}$ , der nicht nur den Nullvektor enthält. Für eine Teilmenge $\mathcal{B} \subset V$ sind dann folgende Aussagen äquivalent:

- (i)  $\mathcal{B}$  ist eine Basis.

- (ii) Jeder Vektor in  $V$  kann sich auf eindeutige Art und Weise als Linearkombination von Vektoren aus  $\mathcal{B}$  schreiben.
- (iii)  $\mathcal{B}$  ist ein minimales Erzeugendensystem, d.h. für jede echte Teilmenge  $C \subset \mathcal{B}$ ,  $C \neq \mathcal{B}$ , gilt, dass  $\text{span}_{\mathbb{K}}(C) \neq V$ .
- (iv)  $\mathcal{B}$  ist eine maximal linear unabhängige Teilmenge von  $V$ , d.h. jede echte Obermenge  $C \supset \mathcal{B}$ ,  $C \neq \mathcal{B}$ , ist linear abhängig.

**Beweis.** Der Beweis wird durch die Teilschritte (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Rightarrow$ (i) geführt.

(i) $\Rightarrow$ (ii): Dies wird durch Negation gezeigt, d.h. wir nehmen an, dass (ii) nicht gilt und folgern, dass auch (i) dann nicht gilt. Wir nehmen also an, dass nicht jeder Vektor in  $v \in V$  auf eindeutige Art und Weise als Linearkombination von Vektoren aus  $\mathcal{B}$  geschrieben werden. Dies kann durch zwei sich nicht ausschließende verletzt werden: es gibt einen Vektor  $v$ , der nicht in  $\text{span}_{\mathbb{K}}(\mathcal{B})$  liegt, oder es gibt einen Vektor  $w$ , der auf verschiedene Art und Weise als Linearkombination von Vektoren aus  $\mathcal{B}$  geschrieben werden kann. In ersterem Fall ist die Erzeugendeneigenschaft von  $\mathcal{B}$  verletzt, d.h.  $\mathcal{B}$  ist keine Basis. In zweiterem Fall seien also zwei verschiedene Linearkombinationen gegeben:

$$w = \sum_{j=1}^J \lambda_j b_j = \sum_{j=1}^J \mu_j b_j,$$

wobei  $J \in \mathbb{N}$ ,  $b_j \in \mathcal{B}$  und  $\lambda_j, \mu_j \in \mathbb{K}$  (beachte, dass wenn die in den beiden Darstellungen auftretenden  $b_j$ 's zunächst nicht gleich sind, so kann die Vereinigung von beiden betrachtet werden, indem die entsprechenden Koeffizienten  $\lambda_j$  und/oder  $\mu_j$  gleich 0 gesetzt werden). Nun folgt aber durch Umschreiben

$$\sum_{j=1}^J (\lambda_j - \mu_j) b_j = 0,$$

was eine nicht-triviale Linearkombination des Nullvektors ist, da ja nicht  $\lambda_j = \mu_j$  für alle  $j = 1, \dots, J$ . Somit ist  $\mathcal{B}$  nicht linear unabhängig und  $\mathcal{B}$  somit wiederum keine Basis.

(ii) $\Rightarrow$ (iii): Aus (ii) folgt offensichtlich, dass  $\mathcal{B}$  ein Erzeugendensystem ist. Es verbleibt noch zu zeigen, dass es minimal ist. Sei also  $C \subset \mathcal{B}$  eine echte Teilmenge. Wir nehmen an, dass  $\text{span}_{\mathbb{K}}(C) = V$ . Dann gibt es ein  $b \in \mathcal{B} \setminus C$  und dieser Vektor  $b$  lässt sich als Linearkombination von Vektoren aus  $C$  darstellen. Andererseits gibt es die Linearkombination  $b = 1 \cdot b$ , im Widerspruch zu der eindeutigen Darstellbarkeit. Somit kann  $\text{span}_{\mathbb{K}}(C) = V$  nicht gelten.

(iii) $\Rightarrow$ (iv): Zunächst wird gezeigt, dass  $\mathcal{B}$ , mit den Eigenschaften aus (iii), linear unabhängig ist, wiederum mit einem Widerspruchsbeweis. Sei also  $\mathcal{B}$  linear abhängig. Dann gibt es einen Vektor  $w \in \mathcal{B}$ , der Linearkombination der anderen Vektoren aus  $\mathcal{B}$  ist. Sei  $C = \mathcal{B} \setminus \{w\}$ , was eine echte Teilmenge von  $\mathcal{B}$  ist. Aber  $\text{span}_{\mathbb{K}}(C) = \text{span}_{\mathbb{K}}(\mathcal{B}) = V$ , im Widerspruch zur Voraussetzung (iii). Somit ist  $\mathcal{B}$  linear unabhängig. Es verbleibt noch zu zeigen, dass  $\mathcal{B}$  maximal als linear unabhängige Menge ist. Sei  $C'$  eine echte Obermenge von  $\mathcal{B}$ , die linear unabhängig ist. Dann gibt es einen Vektor  $w \in C' \setminus \mathcal{B}$ . Da aber  $\mathcal{B}$  ein Erzeugendensystem ist, ist  $w$  Linearkombination von Vektoren aus  $\mathcal{B}$  und somit wäre  $C'$  linear abhängig.

(iv) $\Rightarrow$ (i): Da  $\mathcal{B}$  linear unabhängig ist, bleibt nur noch die Erzeugendeneigenschaft  $\text{span}_{\mathbb{K}}(\mathcal{B}) = V$  nachzuweisen. Nach dem Basiserweiterungssatz 2.17 gibt es eine Basis  $\mathcal{B}'$  mit  $\mathcal{B} \subset \mathcal{B}'$ . Insbesondere ist  $\mathcal{B}'$  linear unabhängig. Wäre  $\mathcal{B}$  echte Teilmenge von  $\mathcal{B}'$ , so wäre die in (iv) gegebene Maximalität verletzt. Also ist  $\mathcal{B} = \mathcal{B}'$  eine Basis.  $\square$

**2.19 Satz** (Austauschsatz von Steinitz) Sei  $\mathcal{B} = \{b_1, \dots, b_N\}$  eine Basis des Vektorraumes  $V$ . Seien  $a_1, \dots, a_k$  linear unabhängige Vektoren in  $V$ . Dann gibt unterschiedliche Indizes  $i_1, \dots, i_{N-k} \in \{1, \dots, N\}$  so dass

$$\mathcal{B}' = \{a_1, \dots, a_k, b_{i_1}, \dots, b_{i_{N-k}}\}$$

eine Basis ist.

**Beweis.** Dies wird durch eine Induktion über  $k$  bewiesen.

Induktionsanfang  $k = 1$ : Es gibt eine eindeutige Zerlegung  $a_1 = \sum_{n=1}^N \lambda_n b_n$ . Sei  $b_i$  so dass  $\lambda_i \neq 0$  (mindestens ein solches  $i$  muss es geben). Nach Umm Nummerierung der Indizes darf  $i = 1$  angenommen werden. Es wird nun nachgewiesen, dass  $\mathcal{B}' = \{a_1, b_2, \dots, b_N\}$  eine Basis ist. Zunächst gilt

$$b_1 = \frac{1}{\lambda_1} \left( a_1 - \sum_{n=2}^N \lambda_n b_n \right). \quad (2.3)$$

Sei nun ein beliebiger Vektor  $v$  gegeben. Da  $\mathcal{B}$  eine Basis ist, gibt es eine eindeutige Linearkombination

$$v = \sum_{n=1}^N \mu_n b_n.$$

Ersetzen von  $b_1$  durch (2.3) ergibt

$$v = \sum_{n=2}^N \mu_n b_n + \frac{\mu_1}{\lambda_1} \left( a_1 - \sum_{n=2}^N \lambda_n b_n \right) = \sum_{n=2}^N \left( \mu_n - \frac{\mu_1 \lambda_n}{\lambda_1} \right) b_n + \frac{\mu_1}{\lambda_1} a_1.$$

Also ist  $v$  im Spann von  $\mathcal{B}'$ , d.h.  $\mathcal{B}'$  ist ein Erzeugendensystem. Es verbleibt noch zu zeigen, dass  $\mathcal{B}'$  linear unabhängig ist. Sei

$$0 = \eta_1 a_1 + \sum_{n=2}^N \eta_n b_n.$$

Einsetzen zeigt

$$0 = \eta_1 \lambda_1 b_1 + \sum_{n=2}^N (\eta_n + \eta_1 \lambda_n) b_n.$$

Da  $\mathcal{B}$  eine Basis ist, folgt  $\eta_1 \lambda_1 = 0$  und  $\eta_n + \eta_1 \lambda_n = 0$  für  $n = 2, \dots, N$ . Da  $\lambda_1 \neq 0$  folgt  $\eta_1 = 0$  und somit auch  $\eta_n = 0$  für  $n = 2, \dots, N$ . Also ist  $\mathcal{B}'$  linear unabhängig.

Induktionsschritt  $k - 1 \rightarrow k$ : Nun sind linear unabhängige  $a_1, \dots, a_k$  gegeben. Nach Induktionsvoraussetzung, gibt es eine geeignete Durchnummerierung der  $b_n$ 's, so dass  $\mathcal{B}' = \{a_1, \dots, a_{k-1}, b_k, \dots, b_N\}$  eine Basis ist. Nun ist  $k - 1 < N$ , denn sonst läge ein Widerspruch zur linearen Unabhängigkeit von  $a_1, \dots, a_k$  vor. Des weiteren ist  $a_k$  eine Linearkombination

$$a_k = \sum_{n=1}^{k-1} \lambda_n a_n + \sum_{n=k}^N \lambda_n b_n.$$

Da  $a_1, \dots, a_k$  linear unabhängig sind und somit  $a_k$  nicht im Spann von  $a_1, \dots, a_{k-1}$  liegt, muss mindestens ein  $\lambda_n$ ,  $n = k, \dots, N$  nicht verschwinden. Nach Umm Nummerierung kann angenommen werden, dass  $\lambda_k \neq 0$ . Analog zu (2.3) erhält man nun

$$b_k = \frac{1}{\lambda_k} \left( a_k - \sum_{n=1}^{k-1} \lambda_n a_n - \sum_{n=k+1}^N \lambda_n b_n \right).$$

Nun kann obiges Argument wiederholt werden um nachzuweisen, dass  $\mathcal{B}'' = \{a_1, \dots, a_k, b_{k+1}, \dots, b_N\}$  eine Basis ist, die aus  $\mathcal{B}'$  durch Ersetzen des Elements  $b_k$  durch  $a_k$  entsteht.  $\square$

**2.20 Korollar** Die Dimension  $\dim_{\mathbb{K}}(V)$  eines Vektorraumes  $V$  über  $\mathbb{K}$  ist wohl-definiert, d.h. unabhängig von der Wahl der Basis.

**Beweis.** Seien  $\mathcal{B} = \{b_1, \dots, b_N\}$  und  $\mathcal{B}' = \{a_1, \dots, a_K\}$  zwei Basen. Nach dem Austauschatz folgt, dass  $K \leq N$ , und aus Symmetriegründen, auch  $N \leq K$ . Somit ist die Dimension  $N = K$ . Besitzt ein Vektorraum eine unendliche Basis, so muss jede Basis unendlich viele Elemente haben.  $\square$

**2.21 Satz** (Dimensionsatz) Seien  $E$  und  $F$  zwei endlich-dimensionale Unterräume eines Vektorraumes  $V$  über  $\mathbb{K}$ . Dann gilt

$$\dim_{\mathbb{K}}(E) + \dim_{\mathbb{K}}(F) = \dim_{\mathbb{K}}(E \cap F) + \dim_{\mathbb{K}}(E + F),$$

wobei  $E + F = \text{span}_{\mathbb{K}}(E \cup F)$ .

**2.22 Beispiel** Sei  $V = \mathbb{R}^3$  und seien  $E$  und  $F$  zwei-dimensionale Unterräume, d.h. zwei Ebenen durch den Ursprung. Also gilt  $\dim_{\mathbb{R}}(E) + \dim_{\mathbb{R}}(F) = 2 + 2 = 4$ . Typischerweise schneiden diese Ebenen sich in einer Gerade, die auch den Ursprung enthält, so dass  $E \cap F$  eindimensional ist und  $E \cup F = \mathbb{R}^3$ . In dieser typischen Situation ist also  $\dim_{\mathbb{R}}(E \cap F) + \dim_{\mathbb{R}}(E \cup F) = 1 + 3 = 4$ , wie im Satz behauptet. Es gibt allerdings auch die untypische Situation (oder oft auch nicht generische Situation genannt), in der die beiden Ebenen gleich sind und somit  $\dim_{\mathbb{R}}(E \cap F) + \dim_{\mathbb{R}}(E \cup F) = 2 + 2 = 4$ , wiederum im Einklang mit dem Satz.  $\diamond$

**Beweis** von Satz 2.21. Zunächst ist  $E \cap F$  ein Unterraum sowohl von  $E$  als auch  $F$ , also ist seine Dimension endlich. Sei  $b_1, \dots, b_N$  eine Basis von  $E \cap F$ , d.h.  $N = \dim_{\mathbb{K}}(E \cap F)$ . Nach dem Basiserweiterungssatz gibt es dann Vektoren  $e_1, \dots, e_K$ , so dass  $b_1, \dots, b_N, e_1, \dots, e_K$  eine Basis von  $E$  ist, und es gibt Vektoren  $f_1, \dots, f_L$ , so dass  $b_1, \dots, b_N, f_1, \dots, f_L$  eine Basis von  $F$  ist. Also  $N + K = \dim_{\mathbb{K}}(E)$  und  $N + L = \dim_{\mathbb{K}}(F)$ . Nun spannen  $b_1, \dots, b_N, e_1, \dots, e_K, f_1, \dots, f_L$  auch  $E + F$  auf. Es verbleibt noch zu zeigen, dass diese  $N + K + L$  Vektoren auch linear unabhängig sind. Sei

$$0 = \sum_{n=1}^N \lambda_n b_n + \sum_{k=1}^K \mu_k e_k + \sum_{l=1}^L \eta_l f_l.$$

Dann ist

$$\sum_{n=1}^N \lambda_n b_n + \sum_{k=1}^K \mu_k e_k = - \sum_{l=1}^L \eta_l f_l \in E \cap F.$$

Hieraus folgt zunächst  $\eta_l = 0$ , da  $f_l \notin E \cap F$ . Analog folgt  $\mu_k = 0$ . Zuletzt liegt dann die Gleichung  $\sum_{n=1}^N \lambda_n b_n = 0$  vor, woraus  $\lambda_n = 0$  folgt, da  $b_n$  eine Basis von  $E \cap F$  bilden. Also bilden die Vektoren  $b_1, \dots, b_N, e_1, \dots, e_K, f_1, \dots, f_L$  eine Basis von  $E + F$  und  $\dim_{\mathbb{K}}(E + F) = N + K + L$ . Hieraus folgt nun die Behauptung.  $\square$

### 3 Lineare Gleichungssysteme und der Gauß-Algorithmus

Weiterhin bezeichnet  $\mathbb{K}$  entweder die reellen Zahlen  $\mathbb{R}$  oder die komplexen Zahlen  $\mathbb{C}$ .

**3.1 Definition** Eine lineare Gleichung in  $M$  sogenannten Unbekannten  $x_1, \dots, x_M \in \mathbb{K}$  ist von der Form

$$a_1x_1 + a_2x_2 + \dots + a_Mx_M = b,$$

wobei  $a_1, \dots, a_M$  die Koeffizienten der Gleichung und  $b \in \mathbb{K}$  die sogenannte rechte Seite oder Inhomogenität ist. Ein lineares Gleichungssystem in  $M$  Unbekannten  $x_1, \dots, x_M \in \mathbb{K}$  ist eine Menge von endlich vielen linearen Gleichungen, sagen wir  $N \in \mathbb{N}$  solcher Gleichungen. Um diese Gleichungen auszuschreiben, werden sogenannte Doppelindizes verwendet:

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,M}x_M &= b_1, \\ a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,M}x_M &= b_2, \\ &\vdots \\ a_{N,1}x_1 + a_{N,2}x_2 + \dots + a_{N,M}x_M &= b_N. \end{aligned} \tag{3.1}$$

Hierbei sind  $a_{n,m} \in \mathbb{K}$  und  $b_n \in \mathbb{K}$  für alle  $m = 1, \dots, M$  und  $n = 1, \dots, N$ . Des weiteren heißt  $n$  der Zeilenindex von  $a_{n,m}$  und  $m$  der Spaltenindex von  $a_{n,m}$ . Gesucht sind dann  $x_1, \dots, x_M \in \mathbb{K}$ , die simultan alle  $N$  Gleichungen erfüllen.

Das lineare Gleichungssystem heißt homogen falls  $b_1 = \dots = b_N = 0$ , ansonsten heißt es inhomogen.

**3.2 Bemerkungen** Linearität einer Gleichung in  $x_1, \dots, x_M$  bedeutet, dass keine höheren Potenzen wie z.B.  $(x_1)^2$  oder  $(x_2)^5$  auftauchen, aber auch keine anderen nicht-linearen Funktionen von  $x_1, \dots, x_M$  wie die trigonometrischen Funktionen oder die Exponentialfunktion. Ein Beispiel für eine nicht-lineare Gleichung ist also  $(x_1)^3 + \sin(x_2) = 1$ . Wir werden sehen, dass es recht einfache Methoden zum Lösen linearer Gleichungen gibt (nämlich den Gauß-Algorithmus). Hingegen ist das Lösen nicht-linearer Gleichungen im Allgemeinen sehr schwierig. Des weiteren sei schon hier bemerkt, dass es möglich ist, dass ein Gleichungssystem wie (3.1) gar keine oder unendlich viele Lösungen haben kann.  $\diamond$

**3.3 Beispiele** Lineare Gleichungssysteme treten in vielen Anwendungen auf natürliche Art und Weise auf, z.B. weil viele Grundgesetze in den Naturwissenschaften durch lineare Gleichungen gegeben sind. In der jetzt folgenden Liste von Beispielen wird aufgezeigt, dass auch schon die wenigen Begriffe, die im Kapitel 2 eingeführt wurden, zu linearen Gleichungssystemen führen.

1. Sei  $V = \mathbb{K}^N$ . Untersucht werden soll, ob  $M$  Vektoren  $v_1, \dots, v_M \in V$  linear abhängig oder unabhängig sind. Also sind gesucht Skalare  $\lambda_1, \dots, \lambda_M \in \mathbb{K}$  mit

$$\sum_{m=1}^M \lambda_m v_m = 0, \tag{3.2}$$

und es soll untersucht werden, ob es mehr als die triviale Lösung  $\lambda_1 = \dots = \lambda_M = 0$  gibt. Diese Gleichung ist in der Tat ein homogenes lineares Gleichungssystem in den Unbekannten  $\lambda_1, \dots, \lambda_M$ , denn wenn

$$v_m = \begin{pmatrix} v_{1,m} \\ \vdots \\ v_{N,m} \end{pmatrix}, \tag{3.3}$$

so ist (3.2) explizit geschrieben

$$\begin{aligned} v_{1,1}\lambda_1 + v_{1,2}\lambda_2 + \dots + v_{1,M}\lambda_M &= 0, \\ v_{2,1}\lambda_1 + v_{2,2}\lambda_2 + \dots + v_{2,M}\lambda_M &= 0, \\ &\vdots \\ v_{N,1}\lambda_1 + v_{N,2}\lambda_2 + \dots + v_{N,M}\lambda_M &= 0. \end{aligned}$$

Also, wenn wir lernen, methodisch die Lösungen linearer Gleichungssysteme zu bestimmen, so können wir auch die Frage nach der linearen Unabhängigkeit in konkreten Situationen beantworten.

2. Wieder sei  $V = \mathbb{K}^N$  und  $v_1, \dots, v_M \in V$  dargestellt wie in (3.3). Des Weiteren sei  $w \in V$  und die Frage ist, ob  $w \in \text{span}(\{v_1, \dots, v_M\})$ . Hierzu sind gesucht  $\lambda_1, \dots, \lambda_M \in \mathbb{K}$ , so dass

$$\sum_{m=1}^M \lambda_m v_m = w = \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix}.$$

Explizit ausgeschrieben ist dies wieder ein lineares Gleichungssystem in  $\lambda_1, \dots, \lambda_M$ , diesmal inhomogen:

$$\begin{aligned} v_{1,1}\lambda_1 + v_{1,2}\lambda_2 + \dots + v_{1,M}\lambda_M &= w_1, \\ v_{2,1}\lambda_1 + v_{2,2}\lambda_2 + \dots + v_{2,M}\lambda_M &= w_2, \\ &\vdots \\ v_{N,1}\lambda_1 + v_{N,2}\lambda_2 + \dots + v_{N,M}\lambda_M &= w_N. \end{aligned}$$

3. Wenn  $b_1, \dots, b_N$  eine Basis des  $\mathbb{K}^N$  ist, so besagt Satz 2.18, dass jeder Vektor  $w \in \mathbb{K}^N$  auf eindeutige Art und Weise als Linearkombination von  $b_1, \dots, b_N$  geschrieben werden kann. Um diese Linearkombination explizit anzugeben, kann wieder genau wie im vorherigen Beispiel vorgegangen werden, nur dass jetzt *a priori* bekannt ist, dass es genau eine Lösung des zugehörigen linearen Gleichungssystems gibt.  $\diamond$

Der Gauß-Algorithmus ist eine Methode, um lineare Gleichungssysteme zu lösen, die auch effizient auf einem Computer implementiert werden kann. Der Algorithmus besteht aus sogenannten elementaren Umformungen:

**3.4 Definition** *Elementare Umformungen eines linearen Gleichungssystems sind:*

- (i) *Multiplikation einer Zeile mit einem nicht verschwindenden Skalar aus  $\mathbb{K}$ .*
- (ii) *Ersetzen einer Zeile durch ihre Summe mit einer anderen Zeile.*
- (iii) *Vertauschen von zwei Zeilen des Gleichungssystems.*

**3.5 Satz** *Elementare Umformungen ändern die Lösungsmenge eines linearen Gleichungssystems nicht.*

**Beweis.** Dies ist offensichtlich für Umformungen (i) und (iii). Bzgl. (ii) ist es so, dass das Addieren von wahren Gleichungen wieder wahre Gleichungen produziert, also wird die Lösungsmenge höchstens größer. Aber die Umformung (ii) ist auch umkehrbar (Addieren der mit  $-1$  multiplizierten Zeile, die vorher addiert wurde), was die Lösungsmenge wieder höchstens größer macht. Zusammen ergibt sich, dass die Lösungsmenge unverändert bleibt.  $\square$

Ziel des Gauß-Algorithmus ist nun die sogenannte Zeilenstufenform zu erreichen, an der die Lösungsmenge dann direkt abgelesen werden kann, wie wir weiter unten sehen werden.

**3.6 Definition** *Ein lineares Gleichungssystem (3.1) ist in der Zeilenstufenform, wenn Folgendes gilt: wenn für die  $m$ te Gleichung  $a_{m,1} = \dots = a_{m,k} = 0$ , dann gilt für die  $(m+1)$ ste Gleichung  $a_{m+1,1} = \dots = a_{m+1,k+1} = 0$ . Die Zeilenstufenform heißt normiert, wenn zudem der erste nicht verschwindende Koeffizient in jeder Zeile gleich 1 ist, wobei die Ordnung in jeder Zeile von dem Spaltenindex her stammt.*

Der Gauß-Algorithmus besteht nun aus elementaren Umformungen in einer wohl bestimmten Reihenfolge, so dass am Ende des Algorithmus eine normierte Zeilenstufenform vorliegt. Zunächst sei dies an einem Beispiel illustriert.

**3.7 Beispiel** Gegeben ist ein lineares Gleichungssystem in drei Variablen  $x_1, x_2, x_3$  bestehend aus drei Gleichungen:

$$\begin{aligned} 2x_1 + x_2 + 3x_3 &= 0 \\ 4x_1 - x_2 + x_3 &= 10 \\ 3x_1 + 2x_2 + x_3 &= 7 \end{aligned}$$

Im ersten Schritt wird die erste Zeile normiert, hier durch Multiplikation mit  $\frac{1}{2}$ :

$$\begin{aligned} x_1 + \frac{1}{2}x_2 + \frac{3}{2}x_3 &= 0 \\ 4x_1 - x_2 + x_3 &= 10 \\ 3x_1 + 2x_2 + x_3 &= 7 \end{aligned}$$

Sollte dies nicht möglich sein (wenn der Eintrag vor  $x_1$  in der ersten Zeile verschwindet), wird zuerst mit der zweiten oder notfalls dritten Zeile vertauscht. Die erste Zeile wird nun im Weiteren nicht mehr verändert. Im zweiten Schritt wird die erste Spalte bearbeitet, indem zur zweiten und dritten Spalte adäquate Vielfache der ersten Gleichung dazu addiert werden:

$$\begin{aligned} x_1 + \frac{1}{2}x_2 + \frac{3}{2}x_3 &= 0 \\ 4x_1 - x_2 + x_3 &= 10 && \text{addierte } (-4) \text{ mal die erste Gleichung} \\ 3x_1 + 2x_2 + x_3 &= 7 && \text{addierte } (-3) \text{ mal die erste Gleichung} \end{aligned}$$

was ergibt

$$\begin{aligned} x_1 + \frac{1}{2}x_2 + \frac{3}{2}x_3 &= 0 \\ -3x_2 - 5x_3 &= 10 \\ \frac{1}{2}x_2 - \frac{7}{2}x_3 &= 7 \end{aligned}$$

Nun wird die zweite Zeile normiert, wiederum falls dies möglich ist, ansonsten wird mit den unteren Zeilen vertauscht (hier wäre nur noch eine Zeile möglich). Dann wird die zweite Spalte bearbeitet, durch Addieren adäquater Vielfacher der zweiten Zeile:

$$\begin{aligned} x_1 + \frac{1}{2}x_2 + \frac{3}{2}x_3 &= 0 \\ x_2 + \frac{5}{3}x_3 &= -\frac{10}{3} \\ \frac{1}{2}x_2 - \frac{7}{2}x_3 &= 7 && \text{addierte } \left(-\frac{1}{2}\right) \text{ mal die zweite Gleichung} \end{aligned}$$

Dies ergibt schon eine Zeilenstufenform:

$$\begin{aligned} x_1 + \frac{1}{2}x_2 + \frac{3}{2}x_3 &= 0 \\ x_2 + \frac{5}{3}x_3 &= -\frac{10}{3} \\ -\frac{26}{6}x_3 &= \frac{26}{3} \end{aligned}$$



die lediglich noch normiert werden muss:

$$\begin{aligned}x_1 + \frac{1}{2}x_2 + \frac{3}{2}x_3 &= 0 \\x_2 - \frac{5}{3}x_3 &= \frac{10}{3} \\x_3 &= -2\end{aligned}$$

Jetzt ist die Lösung sehr einfach abzulesen:  $x_3 = -2$  ist klar, aus der zweiten Gleichung folgt nach Einsetzen  $x_2 = \frac{10}{3} - 2\frac{5}{3} = 0$  und schließlich aus der ersten Gleichung  $x_1 = -2\frac{3}{2} = 3$ .  $\diamond$

Nach diesem Beispiel, können wir nun den Gauß-Algorithmus allgemein beschreiben, und zwar in einer Form, die problemlos auf dem Computer implementiert werden kann.

### Schritte des Gauß-Algorithmus für die normierte Zeilenstufenform:

- (i) Wenn  $a_{1,1} \neq 0$ , gehe zum nächsten Schritt. Falls  $a_{1,1} = 0$ , bestimme das kleinste  $n$  so dass  $a_{n,1} \neq 0$  und vertausche Zeilen 1 und  $n$ . Wenn  $a_{n,1} = 0$  für alle  $n = 1, \dots, N$ , so eliminiere die erste Spalte und somit die erste Variable  $x_1$ , gehe zur zweiten Spalte und gehe analog vor (dies bedeutet, dass die erste Variable  $x_1$  nicht spezifiziert ist und somit jeden Wert in  $\mathbb{K}$  annehmen kann und dies keinen Einfluss auf die anderen Variablen hat). Falls auch die zweite Spalte nur Nuller enthält, eliminiere sie und Variable  $x_2$  auch, etc. Danach liegt ein Gleichungssystem vor (evtl. mit weniger Variablen), welches in der ersten Zeile einen ersten nicht-verschwindenden Eintrag hat, der nach Umnummerierung der Indizes wieder mit  $a_{1,1} \neq 0$  bezeichnet wird.
- (ii) Normiere die erste Zeile, indem sie durch  $a_{1,1}$  geteilt wird.
- (iii) Für alle verbleibenden Zeilen wird Folgendes durchgeführt: zur  $n$ ten Zeile wird das  $(-a_{n,1})$ -fache der ersten (normierten) Zeile hinzuaddiert. Dann enthält die erste Spalte nur eine 1 und sonst Nullen.
- (iv) Betrachte nun das kleinere Gleichungssystem, was durch Vernachlässigen der ersten Zeile und ersten Spalte entsteht. Beginne wieder mit Schritt (i) für dieses kleinere Gleichungssystem, es sei denn es ist nur noch eine Zeile vorhanden, die dann lediglich normiert wird.

### Ablezen der Lösungsmenge von der normierten Zeilenstufenform:

- (i) Wenn es eine Zeile (Gleichung) gibt, in der keine Unbekannten mehr auftreten, aber die Inhomogenität nicht verschwindet, so hat das Gleichungssystem keine Lösung.
- (ii) Anderenfalls enthält die unterste nicht-verschwindende Zeile die Unbekannten  $x_{M-r}, \dots, x_M$ , wobei  $r \geq 0$ . Somit, wenn es die  $n$ te Zeile ist,

$$x_{M-r} + a_{n,M-r+1}x_{M-r+1} + \dots + a_{n,M}x_M = b_n. \tag{3.4}$$

Falls  $r = 0$ , so ist  $x_M = b_n$  eindeutig bestimmt. Falls  $r \geq 1$ , so gibt es  $r$  freie Parameter  $x_M = t_1 \in \mathbb{K}, \dots, x_{M-r+1} = t_r \in \mathbb{K}$ . Dann ist  $x_{M-r}$  festgelegt.

(iii) Nun gehe iterativ zur jeweils vorherigen Zeile  $n - 1$  vor und ersetze alle vorher schon berechneten und durch freie Parameter ersetzten Variablen. Wiederum verbleibt ein Gleichung der Art (3.4), die dann genauso zu Lösungen und evtl. freien Parametern führt wie im Schritt (ii).

Das Bestimmen der Lösungsmenge sei nun an einem Beispiel erläutert:

**3.8 Beispiel** Gegeben sei das reelle lineare Gleichungssystem

$$\begin{aligned}x_1 + 2x_2 + 3x_3 - x_4 &= 0 \\x_3 + x_4 &= 2\end{aligned}$$

Dieses System ist schon auf normierter Zeilenstufenform, so dass die Lösung direkt abgelesen werden kann. Sei  $x_4 = t_1 \in \mathbb{R}$ . Dann ist  $x_3 = 2 - t_1$ . Einsetzen in die erste Gleichung ergibt:

$$x_1 + 2x_2 = -3(2 - t_1) + t_1 = -6 + 4t_1.$$

Also wird  $x_2 = t_2 \in \mathbb{R}$  als zweiter freier Parameter gewählt und dann ist  $x_1 = -6 + 4t_1 - 2t_2$ . Also ist die Lösungsmenge

$$\mathbb{L} = \left\{ \left( \begin{array}{c} -6 + 4t_1 - 2t_2 \\ t_2 \\ 2 - t_1 \\ t_1 \end{array} \right) \middle| t_1, t_2 \in \mathbb{R} \right\} = \left\{ \left( \begin{array}{c} -6 \\ 0 \\ 2 \\ 0 \end{array} \right) + t_1 \left( \begin{array}{c} 4 \\ 0 \\ -1 \\ 1 \end{array} \right) + t_2 \left( \begin{array}{c} -2 \\ 1 \\ 0 \\ 0 \end{array} \right) \middle| t_1, t_2 \in \mathbb{R} \right\}.$$

In der zweiten Darstellung sind die beiden Vektoren, die mit  $t_1$  und  $t_2$  skalarmultipliziert werden, die sogenannten Richtungsvektoren und der verbleibende konstante Vektor der sogenannte Richtungsvektor.

Keiner dieser Vektoren ist eindeutig bestimmt, d.h. alle können abgeändert werden ohne die Lösungsmenge zu ändern, es sei denn die Lösung ist durch einen Punkt gegeben. Um dies zu illustrieren, wählen wir  $x_3 = s_1 \in \mathbb{R}$  und  $x_2 = s_2 \in \mathbb{R}$  als freie Parameter. Dann ergibt sich  $x_4 = 2 - s_1$  und  $x_1 = 2 - 4s_1 - 2s_2$  und somit

$$\mathbb{L} = \left\{ \left( \begin{array}{c} 2 - 4s_1 - 2s_2 \\ s_2 \\ s_1 \\ 2 - s_1 \end{array} \right) \middle| s_1, s_2 \in \mathbb{R} \right\} = \left\{ \left( \begin{array}{c} 2 \\ 0 \\ 0 \\ 2 \end{array} \right) + s_1 \left( \begin{array}{c} -4 \\ 0 \\ 1 \\ -1 \end{array} \right) + s_2 \left( \begin{array}{c} -2 \\ 1 \\ 0 \\ 0 \end{array} \right) \middle| s_1, s_2 \in \mathbb{R} \right\}.$$

Obwohl dies anders aussieht, liegt die gleiche Menge vor. Offensichtlich geht die zweite in die erste Darstellung über, wenn man  $t_1 = 2 - s_1$  und  $t_2 = s_2$  setzt.  $\diamond$

Bevor die Struktur der Lösungsmenge weiter erläutert wird, sollen nun zunächst Matrizen und die Matrix-Vektor-Multiplikation eingeführt werden.

**3.9 Definition** Gegeben seien Skalare  $A_{n,m} \in \mathbb{K}$  für Indizes  $n = 1, \dots, N$  und  $m = 1, \dots, M$  (genau wie in dem linearen Gleichungssystem (3.1)). Die zugehörige  $N \times M$  Matrix  $A$  ist folgende Anordnung dieser Zahlen:

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,M} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,M} \\ \vdots & & & \vdots \\ A_{N,1} & A_{N,2} & \cdots & A_{N,M} \end{pmatrix}. \quad (3.5)$$

Es wird auch folgende Kurzschreibweise verwandt:

$$A = (A_{n,m})_{n=1,\dots,N, m=1,\dots,M}.$$

Eine  $M \times 1$  Matrix ist dann ein Vektor im  $\mathbb{K}^M$  und anstelle  $x = (x_{m,k})_{m=1,\dots,M, k=1} \in \mathbb{K}^M$  wird auch geschrieben

$$x = (x_m)_{m=1,\dots,M} = \begin{pmatrix} x_1 \\ \vdots \\ x_M \end{pmatrix}. \quad (3.6)$$

Für eine  $N \times M$  Matrix  $A$  und eine  $M \times K$  Matrix  $B = (B_{m,k})_{m=1,\dots,M, k=1,\dots,K}$  wird nun das Matrixprodukt  $AB$  als folgende  $N \times K$  Matrix definiert:

$$AB = \left( \sum_{m=1}^M a_{n,m} B_{m,k} \right)_{n=1,\dots,N, k=1,\dots,K} = (A_{n,1}B_{1,k} + A_{n,2}B_{2,k} + \dots + A_{n,M}B_{M,k})_{n=1,\dots,N, k=1,\dots,K}. \quad (3.7)$$

Wenn  $A$  eine  $N \times M$  Matrix und  $x \in \mathbb{K}^M$  ist, so ist  $Ax \in \mathbb{K}^N$  wohldefiniert als Matrixprodukt, heißt aber auch Matrix-Vektor-Produkt.

Die Menge der  $N \times M$  Matrizen mit Einträgen in  $\mathbb{K}$  wird mit  $\text{Mat}(N \times M, \mathbb{K})$  bezeichnet.

**Achtung:** Das Matrixprodukt  $AB$  ist nur wohldefiniert, wenn die Anzahl der Spalten von  $A$  gleich der Anzahl der Zeilen von  $B$  ist. In der Definition des Matrixproduktes wird dann über diese Anzahl summiert. Man sagt auch: es wird über den zugehörigen Index kontrahiert.

**3.10 Beispiele** 1. Das Produkt einer  $3 \times 2$  und einer  $2 \times 4$  Matrix ist eine  $3 \times 4$  Matrix:

$$\begin{pmatrix} 1 & 2 \\ i & 3i \\ -i & 0 \end{pmatrix} \begin{pmatrix} 0 & 2 & 3 & 0 \\ i & 0 & -i & 1 \end{pmatrix} = \begin{pmatrix} 2i & 2 & 3 - 2i & 2 \\ -3 & 2i & 3 + 3i & 3i \\ 0 & -2i & -3i & 0 \end{pmatrix}.$$

2. Das Matrixprodukt ist im Allgemeinen nicht kommutativ. Dies sei an einem einfachen Beispiel von  $2 \times 2$  Matrizen gezeigt:

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}.$$

Dann sind

$$AB = \begin{pmatrix} -1 & 4 \\ -1 & 2 \end{pmatrix}, \quad BA = \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix},$$

offensichtlich unterschiedlich. ◇

**3.11 Satz** Das lineare Gleichungssystem (3.1) kann mit Hilfe des Matrix-Vektor-Produktes kompakt geschrieben werden als

$$Ax = b,$$

wobei  $A = (a_{n,m})_{n=1,\dots,N, m=1,\dots,M} \in \text{Mat}(N \times M, \mathbb{K})$ ,  $x \in \mathbb{K}^M$  durch (3.6) gegeben ist und  $b = (b_n)_{n=1,\dots,N} \in \mathbb{K}^N$  die Inhomogenität ist.

**Beweis.** Dies ist lediglich ein Umschreiben von (3.1) mit der Definition des Matrix-Vektor-Produktes. □

Die Matrixschreibweise führt auch auf eine kompaktere Schreibweise für den Gauß-Algorithmus. Dies sei wieder an einem Beispiel illustriert.

**3.12 Beispiel** Wir betrachten das gleiche Beispiel aus 3.7:

$$\begin{aligned} 2x_1 + x_2 + 3x_3 &= 0 \\ 4x_1 - x_2 + x_3 &= 10 \\ 3x_1 + 2x_2 + x_3 &= 7 \end{aligned}$$

Dies wird nun als  $3 \times 4$  Matrix geschrieben, wobei jedoch die 4te Spalte notationell abgetrennt wird und die Matrix-Klammern meist weggelassen werden:

$$\begin{array}{ccc|c} 2 & 1 & 3 & 0 \\ 4 & -1 & 3 & 10 \\ 3 & 2 & 1 & 7 \end{array}$$

Nun sehen die im Gauß-Algorithmus durchgeführten Operationen wie folgt aus. Zunächst wird die erste Zeile normiert:

$$\begin{array}{ccc|c} 1 & \frac{1}{2} & \frac{3}{2} & 0 \\ 4 & -1 & 3 & 10 \\ 3 & 2 & 1 & 7 \end{array} \quad \begin{array}{l} \text{addiere } (-4) \text{ mal die 1. Zeile} \\ \text{addiere } (-3) \text{ mal die 1. Zeile} \end{array}$$

Nun werden die Nullen in der ersten Spalte produziert:

$$\begin{array}{ccc|c} 1 & \frac{1}{2} & \frac{3}{2} & 0 \\ 0 & -3 & -5 & 10 \\ 0 & \frac{1}{2} & -\frac{7}{2} & 7 \end{array}$$

anschließend die zweite Zeile normiert:

$$\begin{array}{ccc|c} 1 & \frac{1}{2} & \frac{3}{2} & 0 \\ 0 & 1 & \frac{5}{3} & -\frac{10}{3} \\ 0 & \frac{1}{2} & -\frac{7}{2} & 7 \end{array} \quad \text{addiere } \left(-\frac{1}{2}\right) \text{ mal die 2. Zeile}$$

und zuletzt die letzte Null produziert

$$\begin{array}{ccc|c} 1 & \frac{1}{2} & \frac{3}{2} & 0 \\ 0 & 1 & \frac{5}{3} & -\frac{10}{3} \\ 0 & 0 & -\frac{26}{6} & \frac{26}{3} \end{array}$$

und normiert:

$$\begin{array}{ccc|c} 1 & \frac{1}{2} & \frac{3}{2} & 0 \\ 0 & 1 & \frac{5}{3} & -\frac{10}{3} \\ 0 & 0 & 1 & -2 \end{array}$$

Beim Ablesen der Lösung können nun  $x_1, x_2, x_3$  wieder eingefügt werden, falls nötig. ◇

**3.13 Definition** Seien  $A = (A_{n,m})_{n=1,\dots,N, m=1,\dots,M}$  und  $B = (B_{n,m})_{n=1,\dots,N, m=1,\dots,M}$  zwei  $N \times M$  Matrizen und  $\lambda \in \mathbb{K}$ . Dann werden zwei neue  $N \times M$  Matrizen definiert durch

$$\lambda A = (\lambda A_{n,m})_{n=1,\dots,N, m=1,\dots,M}, \quad A + B = (A_{n,m} + B_{n,m})_{n=1,\dots,N, m=1,\dots,M}. \quad (3.8)$$

**3.14 Beispiel** Eine Linearkombination zweier  $3 \times 2$  Matrizen ist wieder eine  $3 \times 2$  Matrix:

$$\begin{pmatrix} 1 & 2 \\ i & 3i \\ -i & 0 \end{pmatrix} + 2i \begin{pmatrix} 0 & 2 \\ i & 0 \\ -i & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2+4i \\ -2+i & 3i \\ 2-i & 2i \end{pmatrix}.$$

◇

**3.15 Bemerkung** Wenn alle Spalten einer Matrix  $A \in \text{Mat}(N \times M, \mathbb{K})$  übereinander zu einem Vektor im  $\widetilde{A} \in \mathbb{K}^{NM}$  geschrieben werden, und analog für  $B \in \text{Mat}(N \times M, \mathbb{K})$  ein Vektor  $\widetilde{B} \in \mathbb{K}^{NM}$  eingeführt wird, dann entsprechen die in Definition 3.13 eingeführten Linearkombinationen genau denen für Vektoren im  $\mathbb{K}^{NM}$ , d.h.

$$\widetilde{A + \lambda B} = \widetilde{A} + \lambda \widetilde{B} .$$

Dies folgt, weil die Addition und Skalarmultiplikation jeweils komponentenweise definiert sind. Somit ist der erste Teil (i) folgenden Satzes nicht verwunderlich.  $\diamond$

**3.16 Satz** *Es gelten folgende strukturelle Eigenschaften für die Menge der Matrizen und für die Matrix-Vektor-Multiplikation.*

- (i)  $\text{Mat}(N \times M, \mathbb{K})$  ist ein Vektorraum über  $\mathbb{K}$ , wobei der Nullvektor durch die sogenannte Nullmatrix  $\mathbf{0} = (0)_{n=1, \dots, N, m=1, \dots, M}$  gegeben ist, die nur Nuller enthält.
- (ii)  $\text{Mat}(N \times N, \mathbb{K})$  ist ein Ring mit Eins  $\mathbf{1}$  gegeben durch die sogenannte Einheitsmatrix, nämlich die Diagonalmatrix mit nur Einsen auf der Diagonale:

$$\mathbf{1} = \begin{pmatrix} 1 & 0 & & & \\ 0 & 1 & & & \\ & & \ddots & & \\ & & & 0 & 1 & 0 \\ & & & & & 0 & 1 \end{pmatrix} = (\delta_{n,m})_{n,m=1, \dots, N} ,$$

wobei das Kronecker Delta definiert ist durch

$$\delta_{n,m} = \begin{cases} 1, & n = m, \\ 0, & n \neq m. \end{cases}$$

- (iii) (Linearität der Matrix-Vektor-Multiplikation) Für  $A \in \text{Mat}(N \times M, \mathbb{K})$ ,  $\lambda \in \mathbb{K}$  und  $x, y \in \mathbb{K}^M$ ,

$$A(x + \lambda y) = Ax + \lambda Ay .$$

- (iv) Für  $A \in \text{Mat}(N \times M, \mathbb{K})$ ,  $\lambda \in \mathbb{K}$  und  $B, C \in \text{Mat}(M \times K, \mathbb{K})$ ,

$$A(B + \lambda C) = AB + \lambda AC .$$

- (v) Für  $A, B \in \text{Mat}(N \times M, \mathbb{K})$ ,  $\lambda \in \mathbb{K}$  und  $C \in \text{Mat}(M \times K, \mathbb{K})$ ,

$$(A + \lambda B)C = AC + \lambda BC .$$

- (vi) Für  $A \in \text{Mat}(N \times M, \mathbb{K})$ ,  $B \in \text{Mat}(M \times K, \mathbb{K})$  und  $C \in \text{Mat}(K \times L, \mathbb{K})$ ,

$$(AB)C = A(BC) .$$

**Beweis.** (i) ist nach der Definition offensichtlich, und wurde ja auch schon in Bemerkung 3.15 erläutert. Für (ii) müssen gemäß Definition 1.13 noch die Assoziativität der Matrixmultiplikation sowie die Distributivgesetze nachgewiesen werden. Ersteres ist (vi) und zweitens (iv) und (v), jeweils im Fall  $N = M = K = L$ . All diese Punkte werden gleich etwas allgemeiner nachgewiesen. Außerdem ist zu zeigen, dass  $\mathbf{1}A = A$

gilt, was aber nach (3.7) offensichtlich ist, da jeweils nur ein Summand auftritt. Des Weiteren ist (iii) ein Spezialfall von (iv). Somit sind (iv), (v) und (vi) nachzurechnen. Beginnen wir mit (iv). Seien

$$A = (A_{n,m})_{n=1,\dots,N, m=1,\dots,M}, \quad B = (B_{m,k})_{m=1,\dots,M, k=1,\dots,K}, \quad C = (C_{m,k})_{m=1,\dots,M, k=1,\dots,K},$$

und  $\lambda \in \mathbb{K}$ . Nun wird der Eintrag in der  $n$ ten Zeile und  $k$ ten Spalte von  $A(B + \lambda C)$  gemäß (3.7) und (3.8) berechnet:

$$\begin{aligned} (A(B + \lambda C))_{n,k} &= \sum_{m=1}^M A_{n,m}(B + \lambda C)_{m,k} \\ &= \sum_{m=1}^M A_{n,m}(B_{m,k} + \lambda C_{m,k}) \\ &= \left( \sum_{m=1}^M A_{n,m}B_{m,k} \right) + \lambda \left( \sum_{m=1}^M A_{n,m}C_{m,k} \right) \\ &= (AB)_{n,k} + \lambda(AC)_{n,k}. \end{aligned}$$

Da dies für alle  $n$  und  $k$  gilt, folgt Punkt (iv). Punkt (v) geht analog und für (vi) seien  $A$  und  $B$  wie oben, aber  $C = (C_{k,l})_{k=1,\dots,K, l=1,\dots,L}$ . Dann gilt für alle  $n = 1, \dots, N$  und  $l = 1, \dots, L$ ,

$$\begin{aligned} ((AB)C)_{n,l} &= \sum_{k=1}^K (AB)_{n,k}C_{k,l} \\ &= \sum_{k=1}^K \left( \sum_{m=1}^M A_{n,m}B_{m,k} \right) C_{k,l} \quad (\text{Klammern können weggelassen werden}) \\ &= \sum_{m=1}^M \sum_{k=1}^K A_{n,m}B_{m,k}C_{k,l} \quad (\text{Summationsindize können vertauscht werden}) \\ &= \sum_{m=1}^M A_{n,m} \sum_{k=1}^K B_{m,k}C_{k,l} \quad (k\text{-unabhängiger Faktor } A_{n,m} \text{ aus Summe vorziehen}) \\ &= \sum_{m=1}^M A_{n,m}(BC)_{m,l} \\ &= (A(BC))_{n,l}. \end{aligned}$$

Der Beweis ist beendet. □

Nun liegen alle Hilfsmittel vor, um den Struktursatz für die Lösungsmenge eines linearen Gleichungssystems auch strukturell zu beweisen. Das Gleichungssystem sei

$$Ax = b,$$

wobei  $A \in \text{Mat}(N \times M, \mathbb{K})$  und  $b \in \mathbb{K}^M$ . Gesucht sind also alle  $x \in \mathbb{K}^N$ , die die Gleichung erfüllen, und all diese Lösungen bilden die Lösungsmenge:

$$\mathbb{L}_b = \{x \in \mathbb{K}^N \mid Ax = b\}.$$

### 3.17 Satz (Struktursatz für Lösungen linearer Gleichungssysteme)

- (i) Die Lösungsmenge  $\mathbb{L}_0$  eines homogenen linearen Gleichungssystems ist ein Unterraum des  $\mathbb{K}^N$ .
- (ii)  $\dim_{\mathbb{K}}(\mathbb{L}_0)$  ist die Anzahl der freien Parameter, die beim Ablesen der Lösung von der Zeilenstufenform verwandt werden. Wenn  $K$  die Anzahl der nicht-verschwindenden Zeilen in der Zeilenstufenform ist, dann ist auch  $\dim_{\mathbb{K}}(\mathbb{L}_0) = M - K$ .
- (iii) Die Lösungsmenge  $\mathbb{L}_b$  eines inhomogenen linearen Gleichungssystems ist gegeben durch

$$\mathbb{L}_b = x_0 + \mathbb{L}_0,$$

wobei  $x_0$  eine Lösung des inhomogenen linearen Gleichungssystems ist, auch genannt eine spezielle Lösung.

**Beweis.** (i) Wenn  $x, y \in \mathbb{L}_0$  und  $\lambda \in \mathbb{K}$ , dann gilt nach Satz 3.16(iii)

$$A(x + \lambda y) = Ax + \lambda Ay = 0 + \lambda 0 = 0.$$

Da offensichtlich auch  $0 \in \mathbb{L}_0$ , folgt schon die Behauptung. (ii) Nach Umordnen der Spalten (und somit der Variablen) ist die Zeilenstufenform so, dass  $a_{1,1} = a_{2,2} = \dots = a_{K,K} = 1$ . Die letzte Zeile ist dann von der Gestalt (3.4), d.h. explizit

$$x_K + a_{K,K+1}x_{K+1} + \dots + a_{K,M}x_M = b_K.$$

Nun sind die  $M - K$  freien Parameter  $x_{K+1}, \dots, x_M \in \mathbb{K}$ , und alle anderen Variablen  $x_1, \dots, x_K$  sind durch die freien Parameter ausgedrückt. Somit sind Vektoren des Kernes von der Form

$$\sum_{k=K+1}^M x_k v_k,$$

mit  $v_k - e_k \in \mathbb{K}^M$  hat nicht-verschwindende Einträge nur in den ersten  $K$  Zeilen. Somit sind die Vektoren  $\{v_{K+1}, \dots, v_M\}$  linear unabhängig und, da sie den Kern auch aufspannen (also eine Basis bilden), ist die Dimension des Kernes also  $M - K$ . (iii) Sei  $x_0 \in \mathbb{L}_b$ , d.h.  $Lx_0 = b$  und  $u \in \mathbb{L}_0$ , dann gilt für  $\lambda \in \mathbb{K}$

$$A(x_0 + \lambda u) = Ax_0 + \lambda Au = b + \lambda 0 = b.$$

Somit gilt  $x_0 + \mathbb{L}_0 \subset \mathbb{L}_b$ . Umgekehrt, wenn  $x \in \mathbb{L}_b$ , dann gilt  $A(x - x_0) = b - b = 0$ , d.h.  $x - x_0 \in \mathbb{L}_0$  oder  $x \in x_0 + \mathbb{L}_0$ .  $\square$

Mit folgendem neuen Begriff kann Punkt (ii) von Satz 3.17 auch anders formuliert werden.

**3.18 Definition** Sei  $V$  ein Vektorraum über  $\mathbb{K}$ . Ein affiner Unterraum von  $V$  ist dann eine Teilmenge  $L \subset V$  der Gestalt

$$L = x_0 + U,$$

wobei  $x_0 \in V$  und  $U$  ein Unterraum von  $V$  ist. Der Vektor  $x_0$  heißt auch ein Ortsvektor des Unterraumes. Die Dimension des affinen Unterraumes ist gegeben durch die Dimension von  $U$ .

**3.19 Bemerkungen** Im  $\mathbb{R}^3$  sind die zwei-dimensionalen affinen Unterräume genau die Ebenen (nicht notwendigerweise durch den Ursprung). Dies ist auch ein gutes geometrisches Bild in höheren Dimensionen und bei anderen Körpern. Der Ortsvektor von  $L$  ist nicht eindeutig bestimmt, denn wenn  $x_0$  ein Ortsvektor ist, dann auch  $x_0 + u$  für jedes  $u \in U$ . Des Weiteren ist zu beachten, dass ein affiner Unterraum im Allgemeinen kein Unterraum (mit zusätzlichen Eigenschaften) ist, denn er muss ja nicht den Nullvektor enthalten. Andererseits ist jeder Unterraum auch ein affiner Unterraum. In den Begriffsbildungen der Mathematik ist es leider so, dass ein *rosaroter Panther* weder *rosarot* sein muss, noch eine *Panther*.  $\diamond$

**3.20 Korollar** Die Lösungsmenge eines inhomogenen linearen Gleichungssystems ist ein affiner Unterraum.

## 4 Lineare Abbildungen und ihre darstellenden Matrizen

**4.1 Definition** Seien  $V$  und  $W$  Vektorräume über  $\mathbb{K}$ . Eine Abbildung  $T : V \rightarrow W$  heißt linear genau dann, wenn

$$T(v + \lambda v') = Tv + \lambda T v', \quad \text{für alle } v, v' \in V, \lambda \in \mathbb{K}.$$

Eine lineare Abbildung wird oft auch als linearer Operator bezeichnet.

**4.2 Beispiele** 1. Sei  $V = \mathbb{K}^M$  und  $W = \mathbb{K}^N$ . Dann kann eine Matrix  $A \in \text{Mat}(N \times M, \mathbb{K})$  als lineare Abbildung  $T_A : V \rightarrow W$  aufgefaßt werden:

$$T_A v = Av, \quad v \in V,$$

wobei auf der rechten Seite das Matrix-Vektor-Produkt steht. Dann besagt Satz 3.16(iii) genau, dass  $T_A$  eine lineare Abbildung ist. Meist wird einfach  $T_A = A$  geschrieben, d.h. die Matrix  $A$  wird direkt also lineare Abbildung angesehen. Dieses Beispiel ist in dem Sinne das *Standardbeispiel* für eine lineare Abbildung, das wir später sehen werden, dass jede lineare Abbildung durch eine Matrix dargestellt werden kann, wenn nur Basen von  $V$  und  $W$  gewählt werden.

2. Sei  $V = \text{Mat}(M \times K, \mathbb{K})$  und  $W = \text{Mat}(N \times K, \mathbb{K})$ . Dies sind in der Tat beides Vektorräume gemäß Satz 3.16(i). Wieder definiert  $A \in \text{Mat}(N \times M, \mathbb{K})$  eine lineare Abbildung  $T_A : V \rightarrow W$  durch

$$T_A B = AB, \quad B \in V.$$

Die Linearität von  $T_A$  ist genau Satz 3.16(iv). Wenn  $K = 1$ , ist dieses Beispiel genau das vorherige. Die Abbildung  $T_A$  ist die Linksmultiplikation mit  $A$ . Es ist auch möglich, die Rechtsmultiplikation zu verwenden.

3. Sei  $V = \text{Mat}(K \times N, \mathbb{K})$  und  $W = \text{Mat}(K \times M, \mathbb{K})$ . Jetzt definiert  $A \in \text{Mat}(N \times M, \mathbb{K})$  eine lineare Abbildung  $T_A : V \rightarrow W$  durch

$$T_A B = BA, \quad B \in V.$$

Die Linearität von  $T_A$  ist nun Satz 3.16(v).

4. Nun betrachten wir den unendlich dimensionalen Vektorraum  $V = \mathbb{K}[x]$  der Polynome in der reellen Variable  $x$ . Sei  $a \in \mathbb{K}$  eine Zahl. Zugehörig dazu sei eine Abbildung  $T_a : V \rightarrow \mathbb{K}$  definiert durch

$$T_a(p) = p(a) = \sum_{n=0}^N p_n a^n,$$

wobei  $p(x) = \sum_{n=0}^N p_n x^n$ . Diese Abbildung  $T_a$  heißt *Auswerten bei  $a$*  oder *Einsetzen von  $a$* , und es ist tatsächlich eine lineare Abbildung. Dafür muss überprüft werden, dass für je zwei Polynome  $p, q \in \mathbb{K}[x]$  und jedes Skalar  $\lambda \in \mathbb{K}$  gilt

$$T_a(p + \lambda q) = (p + \lambda q)(a) = p(a) + \lambda q(a) = T_a(p) + \lambda T_a(q).$$

Hierbei folgt das mittlere Gleichungszeichen durch Umordnen der Summe.



5. Wiederum sei  $V = \mathbb{K}[x]$ . Dann definiert das Ableiten (aus der Schule bekannt) eine lineare Abbildung  $T : V \rightarrow V$  durch

$$(Tp)(x) = (\partial_x p)(x).$$

Die Linearität ist die Regel

$$\partial_x(p + \lambda q) = \partial_x p + \lambda \partial_x q, \quad p, q \in \mathbb{K}[x], \quad \lambda \in \mathbb{K},$$

die entweder wieder aus der Schule bekannt ist, oder aber in der Analysis bewiesen wird. Es gibt auch andere Funktionen  $f : \mathbb{R} \rightarrow \mathbb{K}$ , die differenzierbar sind (sagen wir, stetig differenzierbar), ohne ein Polynom zu sein. Die Menge dieser Funktionen wird mit  $C^1(\mathbb{R}, \mathbb{K})$  bezeichnet und  $\partial_x$  ist dann eine lineare Abbildung von  $C^1(\mathbb{R}, \mathbb{K})$  in die Menge  $C(\mathbb{R}, \mathbb{K})$  der stetigen Funktionen. In all diesen unendlich dimensional Beispielen ist das Ableiten eine sogenannte unbeschränkte lineare Abbildung und das Studium solcher linearen Abbildungen ist Bestandteil der Funktionalanalysis.

6. Auch das *Integrieren* oder *Stammfunktionbilden* ist eine lineare Abbildung. Wiederum sei vorausgesetzt, dass aus der Schule bekannt ist, wie Polynome integriert werden, so dass die Abbildung  $T_I : \mathbb{K}[x] \rightarrow \mathbb{K}[x]$  gegeben durch

$$T_I(p) = \sum_{n=0}^N \frac{p_n}{n+1} x^{n+1}, \quad p(x) = \sum_{n=0}^N p_n x^n,$$

tatsächlich als solche erkannt wird (hierbei wurde die Integrationskonstante zu Null gesetzt, was für die Linearität von  $T_I$  notwendig ist).

7. Zuletzt sei  $V = \mathbb{C}^N$  zunächst betrachtet als komplexer Vektorraum. Dann definieren wir die Abbildung *Komplex Konjugieren*  $K : \mathbb{C}^N \rightarrow \mathbb{C}^N$  durch

$$K(x) = \begin{pmatrix} \overline{x_1} \\ \vdots \\ \overline{x_N} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} \in \mathbb{C}^M.$$

Es gilt zwar  $K(x+y) = K(x) + K(y)$ , aber die Abbildung  $K$  ist dennoch nicht komplex linear, weil

$$K(ix) = -i K(x).$$

Andererseits ist die Abbildung reell linear, d.h. linear, wenn  $\mathbb{C}^N$  als reeller Vektorraum aufgefaßt wird. Dann darf nämlich  $i$  nicht als Skalar in der Skalarmultiplikation verwandt werden, und  $K(\lambda x) = \lambda K(x)$  ist richtig für alle  $\lambda \in \mathbb{R}$ .  $\diamond$

**4.3 Definition** Sei  $T : V \rightarrow W$  eine lineare Abbildung zwischen zwei Vektorräumen  $V$  und  $W$  über  $\mathbb{K}$ . Dann ist der Kern von  $T$

$$\text{Ker}(T) = \{x \in V \mid Tx = 0\},$$

und das Bild von  $T$  (auf Englisch 'range') ist

$$\text{Ran}(T) = \{Tx \in W \mid x \in V\}.$$

**4.4 Beispiele** 1. Sei  $A \in \text{Mat}(N \times M, \mathbb{K})$  aufgefaßt als lineare Abbildung  $A : \mathbb{C}^M \rightarrow \mathbb{C}^N$ . Dann ist der Kern

$$\text{Ker}(A) = \{x \in \mathbb{K}^M \mid Ax = 0\},$$

genau die Menge der Lösungen des homogenen linearen Gleichungssystems  $Ax = 0$ . Weiter oben wurde diese Menge mit  $\mathbb{L}_0$  bezeichnet, also  $\text{Ker}(A) = \mathbb{L}_0$ . Mit dem Gauß-Algorithmus haben wir also schon ein Verfahren kennengelernt, den Kern von  $A$  zu berechnen.

2. Oben wurde die Auswertung  $T_a$  bei  $a \in \mathbb{K}$  als lineare Abbildung von  $\mathbb{K}[x]$  nach  $\mathbb{K}$  betrachtet. Dann ist der Kern genau als die Menge der Polynome gegeben, die bei  $a$  eine Nullstelle haben:

$$\text{Ker}(T_a) = \{p \in \mathbb{K}[x] \mid p(a) = 0\} .$$

Andererseits wird jeder Wert in  $\mathbb{K}$  (auf viele verschiedene Arten und Weisen) realisiert als Wert eines Polynomes bei  $a$ . Somit

$$\text{Ran}(T_a) = \mathbb{K} .$$

3. Sei nun  $T = \partial_x$  die Ableitung auf  $\mathbb{K}[x]$ . Dann sind die konstanten Polynome  $p(x) = p_0$  im Kern von  $T$ . Tatsächlich kann gezeigt werden, dass sie genau den Kern bilden. Außerdem ist das Bild  $\text{Ran}(\partial_x) = \mathbb{K}[x]$ , weil nämlich jedes Polynom eine Stammfunktion hat.  $\diamond$

**4.5 Satz** Sei  $T : V \rightarrow W$  eine lineare Abbildung zwischen Vektorräumen  $V, W$  über  $\mathbb{K}$ . Dann:

- (i)  $\text{Ker}(T)$  ist ein Unterraum von  $V$ .
- (ii)  $\text{Ran}(T)$  ist ein Unterraum von  $W$ .
- (iii)  $T$  ist injektiv  $\iff \text{Ker}(T) = \{0\}$
- (iv)  $T$  ist surjektiv  $\iff \text{Ran}(T) = W$

**Beweis.** (i) Zunächst ist  $0 \in \text{Ker}(T)$ , da  $T0 = 0$ . Seien  $x, y \in \text{Ker}(T)$  und  $\lambda \in \mathbb{K}$ . Dann ist  $T(x + \lambda y) = Tx + \lambda Ty = 0 + \lambda 0 = 0$ , so dass auch  $x + \lambda y \in \text{Ker}(T)$  und  $\text{Ker}(T)$  ein Unterraum ist. (ii) Nun seien  $x', y' \in \text{Ran}(T)$ . Dann existieren  $x, y \in V$  mit  $Tx = x'$  und  $Ty = y'$ . Somit  $T(x + \lambda y) = x' + \lambda y' \in \text{Ran}(T)$  und auch  $\text{Ran}(T)$  ist ein Unterraum, da auch  $0 \in \text{Ran}(T)$  wegen  $T0 = 0$ . (iii) Die Hinrichtung ' $\Rightarrow$ ' ist klar, denn bei größerem Kern wäre der Nullvektor mehrfach Bildpunkt. Für die Umkehrung ' $\Leftarrow$ ' sei angenommen, dass  $Tv = Tw$  für zwei unterschiedliche Vektoren  $v, w \in V$ . Dann ist  $T(v - w) = Tv - Tw = 0$ , so dass nicht  $\text{Ker}(T) = \{0\}$  gilt. (iv) ist lediglich die Definition.  $\square$

Es ist auch möglich die Injektivität und Surjektivität einer linearen Abbildung lediglich durch ihr Verhalten auf einer Basis zu untersuchen.

**4.6 Satz** Sei  $T : V \rightarrow W$  eine lineare Abbildung zwischen Vektorräumen  $V, W$  über  $\mathbb{K}$  und  $\mathcal{B}$  eine Basis von  $V$ . Dann:

- (i)  $T$  ist injektiv  $\iff$  Einschränkung  $T|_{\mathcal{B}}$  ist injektiv und  $T(\mathcal{B})$  ist linear unabhängig
- (ii)  $T$  ist surjektiv  $\iff \text{span}(T(\mathcal{B})) = W$
- (iii)  $T$  ist bijektiv  $\iff$  Einschränkung  $T|_{\mathcal{B}}$  ist injektiv und  $T(\mathcal{B})$  ist Basis von  $W$

Falls zudem  $\dim(V) = \dim(W) < \infty$ , so gilt

$$T \text{ ist bijektiv} \iff T(\mathcal{B}) \text{ ist Basis von } W .$$

**Beweis.** Die Basis wird bezeichnet mit  $\mathcal{B} = \{b_1, b_2, \dots\}$ . (i) Für die Hinrichtung ' $\Rightarrow$ ' sei  $\sum_{n=1}^N \lambda_n T b_n = 0$ . Dann ist  $T(\sum_{n=1}^N \lambda_n b_n) = 0$ . Da aber der Kern nach Satz 4.5 trivial ist, folgt  $\sum_{n=1}^N \lambda_n b_n = 0$  und somit  $\lambda_n = 0$  für alle  $n = 1, \dots, N$ . Für ' $\Leftarrow$ ' sei angenommen, dass  $Tv = Tw$  für  $v \neq w$ . Sei  $v - w = \sum_{n=1}^N \lambda_n b_n$  wobei nicht alle  $\lambda_n = 0$  sind. Dann gilt  $0 = T(v - w) = \sum_{n=1}^N \lambda_n T b_n$ . Da alle Bilder  $T b_n$  unterschiedlich sind, folgt, dass  $T(\mathcal{B})$  linear abhängig ist. (ii) Es gilt wegen der Linearität  $\text{span}(T(\mathcal{B})) = T(\text{span}(\mathcal{B})) = T(V)$ , so dass  $\text{Ran}(T) = T(V) = W$  genau dann, wenn  $\text{span}(T(\mathcal{B})) = V$ . (iii) kombiniert (i) und (ii). Für den Zusatz, sei  $\mathcal{B} = \{b_1, \dots, b_N\}$  und somit  $T(\mathcal{B}) = \{T b_1, \dots, T b_N\}$ . Da beides Basen sind und die Dimensionen übereinstimmen, folgt die Injektivität von  $T|_{\mathcal{B}}$ .  $\square$

**4.7 Bemerkung** Es ist nicht möglich, die Bedingung an die Injektivität der Einschränkung  $T|_{\mathcal{B}}$  in (i) und (iii) wegzulassen. Wenn z.B.  $V = \mathbb{R}^3$  und  $W = \mathbb{R}^2$  ist und

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \quad \mathcal{B} = \{e_1, e_2, e_3\},$$

mit den Standardbasisvektoren, dann ist  $T(\mathcal{B}) = \{e_1, e_2\}$  Basis von  $W$ , aber  $T$  ist nicht injektiv, da nämlich  $T(e_2 - e_3) = 0$ , so dass der Kern nicht-trivial ist.  $\diamond$

Nun ist es natürlich sich die Dimensionen der beiden Vektorräume  $\text{Ker}(T)$  und  $\text{Ran}(T)$  anzusehen.

**4.8 Satz** (Dimensionsatz oder Rangsatz) Für jede lineare Abbildung  $T : V \rightarrow W$  gilt

$$\dim(V) = \dim(\text{Ker}(T)) + \dim(\text{Ran}(T)).$$

Die natürlichen Zahlen  $\text{def}(T) = \dim(\text{Ker}(T))$  und  $\text{rk}(T) = \dim(\text{Ran}(T))$  werden auch als der Defekt (auf Englisch “defect” oder manchmal auch “nullity”) und Rang von  $T$  (auf Englisch “rank”) bezeichnet.

**Beweis.** Wenn  $\dim(\text{Ker}(T)) = \infty$ , dann ist auch  $\dim(V) = \infty$  und die Gleichung ist richtig. Sei  $\mathcal{B} = \{b_1, \dots, b_N\}$  eine Basis von dem Vektorraum  $\text{Ker}(T)$ , der ja auch Unterraum von  $V$  ist. Dies wird durch  $\{a_1, \dots, a_M\}$  zu einer Basis  $\mathcal{B}' = \{b_1, \dots, b_N, a_1, \dots, a_M\}$  von  $V$  ergänzt. Dann ist auch  $\{Ta_1, \dots, Ta_M\}$  eine Basis von  $\text{Ran}(T)$ , wobei  $M \in \mathbb{N} \cup \{\infty\}$ . In der Tat, für endliches  $M$  folgt aus

$$0 = \sum_{m=1}^M \lambda_m Ta_m = T \left( \sum_{m=1}^M \lambda_m a_m \right),$$

dass  $\sum_{m=1}^M \lambda_m a_m \in \text{Ker}(T)$ . Da aber  $\{a_1, \dots, a_M\}$  nicht in  $\text{Ker}(T)$  liegen, folgt  $\lambda_1 = \dots = \lambda_M = 0$ . Somit sind  $\{Ta_1, \dots, Ta_M\}$  linear unabhängig (als Teilmenge von  $W$ ). Außerdem kann jeder Vektor  $v \in V$  zerlegt werden bez. der Basis  $\mathcal{B}'$ :

$$v = \sum_{n=1}^N \mu_n b_n + \sum_{m=1}^M \eta_m a_m,$$

und dann impliziert  $Tb_n = 0$  und die Linearität von  $T$

$$Tv = \sum_{m=1}^M \eta_m Ta_m.$$

Dies bedeutet aber

$$\text{Ran}(T) = \text{span}(\{Ta_1, \dots, Ta_M\}).$$

Somit ist also nachgewiesen, dass  $\{Ta_1, \dots, Ta_M\}$  eine Basis von  $\text{Ran}(T)$  ist und deswegen gilt auch  $\dim(\text{Ran}(T)) = M$ . Gleiches gilt für  $M = \infty$ . Da  $\dim(\text{Ker}(T))$  und  $\dim(V) = N + M$  ist die Gleichung bewiesen.  $\square$

**4.9 Bemerkung** In dem Beweis tauchte die Einschränkung  $T'$  der Abbildung  $T$  auf den Unterraum  $V' = \text{span}(\{a_1, \dots, a_M\}) \subset V$  auf. Die Abbildung  $T' : V' \rightarrow W$  hat dann einen trivialen Kern und ist somit injektiv. Der Vektorraum  $V'$  ist eine konkrete Darstellung des sogenannten Faktorraum  $V/\text{Ker}(T)$ , gegeben durch die Faktorgruppe  $V/\text{Ker}(T)$  im Sinne von Satz 1.12. Diese Menge  $V/\text{Ker}(T)$  hat in der hier betrachteten Situation auch eine Vektorraumstruktur. Es ist möglich  $V'$  als ein Komplement zu  $\text{Ker}(T)$  anzusehen in dem Sinne, dass  $V' + \text{Ker}(T) = V$  und  $V' \cap \text{Ker}(T) = \{0\}$ .  $\diamond$

Für den Fall einer linearen Abbildung  $A \in \text{Mat}(N \times M, \mathbb{K})$  kann nun mit dem Gauß-Algorithmus der Kern von  $A$  und somit der Defekt  $\dim(\text{Ker}(A))$  berechnet werden. Daraus kann dann auch der Rang berechnet werden. Allerdings ist es auch möglich, den Rang direkter zu berechnen.

**4.10 Definition** Sei  $A \in \text{Mat}(N \times M, \mathbb{K})$ . Dann ist der Zeilenrang von  $A$  gegeben durch die maximale Anzahl von linear unabhängigen Zeilen von  $A$ , und der Spaltenrang von  $A$  durch die maximale Anzahl von linear unabhängigen Spalten von  $A$ .

**4.11 Satz** Sei  $A \in \text{Mat}(N \times M, \mathbb{K})$ . Dann ist der Rang  $\dim(\text{Ran}(A))$  gleich dem Spaltenrang und gleich dem Zeilenrang. Zudem ist der Rang gleich der Anzahl der nicht verschwindenden Zeilen in der Zeilenstufenform, die aus  $A$  durch elementare Umformungen des Gauß-Algorithmus entsteht.

**Beweis.** Die Spaltenvektoren von  $A$  sind genau die Bilder  $Ae_m$  der Standardbasis  $\{e_1, \dots, e_M\}$  von  $\mathbb{K}^M$ . Zusammen spannen diese Vektoren also  $\text{Ran}(A)$  auf. Die maximale Anzahl von linear unabhängigen Vektoren ist aber gerade die Dimension (siehe Satz 2.18). Somit ist der Spaltenrang von  $A$  gleich dem Rang von  $A$ .

Nun wird noch die Gleichheit von Rang mit Zeilenrang bewiesen, sowie der Zusatz. Hierzu sei zunächst bemerkt, dass die elementaren Umformungen den Zeilenrang nicht verändern (dies ist offensichtlich für das Vertauschen von Zeilen und das Multiplizieren einer Zeile mit einem Skalar, und auch das Ersetzen einer Zeile durch ihre Summe mit einer anderen ändert den Spann der Zeilen und somit den Zeilenrang nicht). Also kann angenommen werden, dass die Matrix schon in Zeilenstufenform ist. Sei  $K$  die Anzahl der nicht verschwindenden Zeilen in der Zeilenstufenform, d.h. der Zeilenrang. An der Zeilenstufenform werden nun die Lösungen des homogenen linearen Gleichungssystems abgelesen und diese Lösungen bilden genau den Kern von  $A$ . Es gibt dann genau  $M - K$  freie Parameter und somit ist die Dimension des Kernes  $M - K$ , siehe Satz 3.17. Der Rangsatz für die lineare Abbildung  $A : \mathbb{K}^M \rightarrow \mathbb{K}^N$  besagt nun, dass

$$K = M - (M - K) = \dim(\mathbb{K}^M) - \dim(\text{Ker}(A)) = \dim(\text{Ran}(A)),$$

was den Beweis beendet. □

**4.12 Beispiel** Sei  $\mu \in \mathbb{R}$  und

$$A = \begin{pmatrix} -1 & \mu & -1 & -1 \\ 4 & 1 & 2 & 1 \\ 2 & 3 & 0 & -1 \end{pmatrix}.$$

Was ist der Rang von  $A$  in Abhängigkeit von  $\mu$ ? Hierzu wird der Gauß-Algorithmus angewandt, um die Matrix auf Zeilenstufenform zu bringen. Zunächst wird die erste Zeile normiert:

$$\begin{pmatrix} 1 & -\mu & 1 & 1 \\ 4 & 1 & 2 & 1 \\ 2 & 3 & 0 & -1 \end{pmatrix}$$

Dann die erste Spalte bearbeitet:

$$\begin{pmatrix} 1 & -\mu & 1 & 1 \\ 0 & 1 + 4\mu & -2 & -3 \\ 0 & 3 + 2\mu & -2 & -3 \end{pmatrix}$$

Nun ist es vorteilhaft (denn es wird vermieden durch  $\mu$ -abhängige Faktoren zu normieren), die zweite und dritte Spalte zu vertauschen, was ja nicht den Spaltenrang verändert.

$$\begin{pmatrix} 1 & 1 & -\mu & 1 \\ 0 & -2 & 1 + 4\mu & -3 \\ 0 & -2 & 3 + 2\mu & -3 \end{pmatrix}$$

Normieren der zweiten Zeile und Bearbeitung der dritten ergibt nun:

$$\begin{pmatrix} 1 & 1 & -\mu & 1 \\ 0 & 1 & -\frac{1}{2} - 2\mu & \frac{3}{2} \\ 0 & 0 & 2 - 2\mu & 0 \end{pmatrix}$$

Somit ist der Rang für  $2 - 2\mu \neq 0$ , d.h.  $\mu \neq 1$ , gleich 3, aber für  $\mu = 1$  ist der Rang lediglich 2.  $\diamond$

Nun folgen einige strukturelle Aussagen über die Menge der linearen Abbildungen.

**4.13 Satz** Seien  $V, W, U$  Vektorräume und  $T, T' : V \rightarrow W$  und  $S : W \rightarrow U$  lineare Abbildungen und  $\lambda \in \mathbb{K}$ . Dann gilt:

(i) Die Hintereinanderausbildung  $S \circ T : V \rightarrow U$  ist eine lineare Abbildung.

(ii) Die Linearkombination  $T + \lambda T' : V \rightarrow W$  definiert durch

$$(T + \lambda T')(v) = T(v) + \lambda T'(v), \quad v \in V, \quad (4.1)$$

ist eine lineare Abbildung.

(iii) Die Menge  $\mathcal{L}(V, W)$  der linearen Abbildungen von  $V$  nach  $W$ , wenn versehen mit der in (4.1) definierten Operation, ist ein Vektorraum. Der Nullvektor ist die Nullabbildung, d.h. die konstante Abbildung mit konstantem Wert  $0(v) = 0$ .

(iv) Die Menge  $\mathcal{L}(V) = \mathcal{L}(V, V)$  der linearen Abbildungen auf  $V$ , d.h. von  $V$  nach  $V$ , bilden mit den in (i) und (ii) definierten Operationen einen Ring  $(\mathcal{L}(V), +, \circ, 0, 1_V)$  mit Eins  $1_V$  definiert durch  $1_V(v) = v$ , d.h.  $1_V$  ist die Identität auf  $V$ . Da in (4.1) zudem noch die Skalarmultiplikation definiert ist, spricht man auch von einer Algebra.

**Beweis.** (i) Seien  $v, v' \in V$  und  $\lambda \in \mathbb{K}$ . Dann

$$S \circ T(v + \lambda v') = S(T(v + \lambda v')) = S(T(v) + \lambda T(v')) = S(T(v)) + \lambda S(T(v')) = S \circ T(v) + \lambda S \circ T(v').$$

(ii) ist eine ähnliche Rechnung, die als Übung durchgeführt werden soll. (iii) und (iv) sind offensichtlich.  $\square$

**4.14 Bemerkung** Alle Aussagen sind Verallgemeinerungen der Aussagen von Satz 3.16. Ein wichtiges Beispiel von Satz 4.13(iv) ist z.B. der Fall  $V = \mathbb{K}^N$ , für den dann  $\mathcal{L}(V) = \text{Mat}(N \times N, \mathbb{K})$  ist. Somit bilden die  $N \times N$  Matrizen  $\text{Mat}(N \times N, \mathbb{K})$  einen Ring und eine Algebra über  $\mathbb{K}$ . Dieser Sachverhalt ist schon in Satz 3.16(ii) bewiesen, so dass Satz 4.13(iv) als Verallgemeinerung davon angesehen werden kann.  $\diamond$

**4.15 Definition** (i) Eine lineare Abbildung  $T \in \mathcal{L}(V, W)$  heißt auch ein Homomorphismus oder Vektorraumhomomorphismus (eine die lineare Struktur erhaltende lineare Abbildung).

(ii) Ein bijektiver Homomorphismus  $T \in \mathcal{L}(V, W)$  heißt auch ein Isomorphismus und in dem Fall, dass ein solcher Isomorphismus existiert, werden die Vektorräume  $V$  und  $W$  auch als isomorph bezeichnet, wofür dann die Notation  $V \cong W$  verwandt wird.

(iii) Eine Abbildung  $T \in \mathcal{L}(V) = \mathcal{L}(V, V)$  heißt auch ein Endomorphismus von  $V$ .

(iv) Eine bijektive Abbildung  $T \in \mathcal{L}(V)$  heißt auch ein Automorphismus von  $V$ .

**4.16 Bemerkung** Nach Satz 4.6(iii) kann ein Isomorphismus wie folgt durch sein Verhalten auf einer Basis  $\mathcal{B}$  von  $V$  charakterisiert werden:

$$T \in \mathcal{L}(V, W) \text{ Isomorphismus} \iff \text{Einschränkung } T|_{\mathcal{B}} \text{ ist injektiv und } T(\mathcal{B}) \text{ ist Basis von } W \quad \diamond$$

Für bijektive Abbildungen kann nun die Umkehrabbildung untersucht werden.

**4.17 Satz** Sei  $T \in \mathcal{L}(V, W)$  bijektiv, d.h. ein Isomorphismus von  $V$  nach  $W$ . Dann ist auch die Umkehrabbildung  $T^{-1} : W \rightarrow V$  linear und wird auch als das Inverse von  $T$  bezeichnet.

**Beweis.** Es ist zu zeigen, dass

$$T^{-1}(v + \lambda w) = T^{-1}(v) + \lambda T^{-1}(w), \quad v, w \in W, \quad \lambda \in \mathbb{K}.$$

Wegen der Bijektivität von  $T$ , ist die Gleichung äquivalent zu

$$T(T^{-1}(v + \lambda w)) = T(T^{-1}(v) + \lambda T^{-1}(w)), \quad v, w \in W, \quad \lambda \in \mathbb{K},$$

was unter Verwendung der Linearität von  $T$  hingegen sofort als wahr erkannt wird. □

**4.18 Korollar** Isomorphie von Vektorräumen ist eine Äquivalenzrelation.

**Beweis.** In der Tat,  $V \cong V$ , denn die Identität ist eine bijektive lineare Abbildung. Somit ist  $\cong$  reflexiv. Auch die Symmetrie  $V \cong W \iff W \cong V$  ist nach Satz 4.17 offensichtlich. Außerdem impliziert  $V \cong W$  und  $W \cong U$  auch  $V \cong U$  (und somit die Transitivität von  $\cong$ ), weil nämlich die Hintereinanderausführung der Isomorphismen gemäß Satz 4.13(i) wieder linear ist, und selbstverständlich auch bijektiv. □

**4.19 Korollar** Sei  $\text{Gl}(V) = \{T \in \mathcal{L}(V) \mid T \text{ bijektiv}\}$ . Dann ist  $\text{Gl}(V)$  versehen mit der Hintereinanderausführung eine Gruppe, genannt die allgemeine lineare Gruppe auf  $V$ . Insbesondere gilt  $TT^{-1} = \mathbf{1}_V$  und  $T^{-1}T = \mathbf{1}_V$  sowie  $(T^{-1})^{-1} = T$ .

**Beweis.** Nach Satz 4.17 ist die bijektive Abbildung  $T^{-1}$  wieder linear, also  $T^{-1} \in \text{Gl}(V)$ . Dass das Rechtsinverse gleich dem Linksinversen ist, folgt nun aus Satz 1.4(i). □

Der folgende Satz erklärt, wie die Umkehrabbildung  $A^{-1}$  zu einer bijektiven linearen Abbildung  $A \in \text{Mat}(N \times N, \mathbb{K})$  berechnet werden kann. Dies Umkehrabbildung  $A^{-1}$  ist wieder durch eine Matrix gegeben, die dann die inverse Matrix heißt.

**4.20 Satz** Sei  $A \in \text{Mat}(N \times N, \mathbb{K})$  bijektiv. Mit Hilfe der elementaren Umformungen des Gauß-Algorithmus wird die Matrix  $(A \mid \mathbf{1}_N) \in \text{Mat}(N \times 2N, \mathbb{K})$  zunächst in die normierte Zeilenstufenform  $(O \mid B) \in \text{Mat}(N \times 2N, \mathbb{K})$  umgeformt, wobei  $O \in \text{Mat}(N \times N, \mathbb{K})$  eine obere Dreiecksmatrix mit Einsen auf der Diagonale ist, dann mit weiteren elementaren Umformungen auf die Form  $(\mathbf{1}_N \mid C) \in \text{Mat}(N \times 2N, \mathbb{K})$ . Dies ist immer möglich und die Matrix  $C \in \text{Mat}(N \times N, \mathbb{K})$  ist dann gleich der inversen Matrix  $C = A^{-1}$ .

**Beweis.** Elementare Umformungen sind Linksmultiplikationen mit bestimmten Matrizen folgender Gestalt:

$$\text{Normieren einer Zeile: } \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & \lambda & \\ & & & \ddots \\ & & & & 1 \end{pmatrix} \quad (\mathbf{1}_N \text{ bis auf eine Eins ersetzt durch } \lambda \neq 0)$$

$$\text{Vertauschen zweier Zeilen: } \begin{pmatrix} 1 & & & \\ & 0 & 1 & \\ & 1 & 0 & \\ & & & \ddots \\ & & & & 1 \end{pmatrix} \quad (\mathbf{1}_N \text{ bis auf zwei Einsen versetzt})$$

$$\text{Addition von Vielfachen einer Zeilen: } \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \ddots & \lambda \\ & & & & \ddots \\ & & & & & 1 \end{pmatrix} \quad (\mathbf{1}_N \text{ mit einem zusätzlichen Eintrag } \lambda)$$

Seien nun  $U(1), \dots, U(k)$  die Matrizen, die bei den elementaren Umformungen bis zur normierten Zeilenstufenform verwandt werden, d.h. mit  $U = U(k) \cdots U(1)$  (beachte Reihenfolge!) gilt  $U(A|\mathbf{1}_N) = (O|B)$ . Hierbei hat  $O$  tatsächlich nur Einsen auf der Diagonale, da der Rang von  $A$  gleich  $N$  ist. Nun wird im sogenannten erweiterten Gauß-Algorithmus zunächst die letzte Spalte von  $O$  durch Addition adäquater Vielfacher der letzten Zeile auf die gewünschte Form gebracht mit einer Eins und sonst nur Nullen, dann die vorletzte Spalte, etc. Seien  $U(k+1), \dots, U(K)$  die Matrizen der dazu notwendigen elementaren Umformungen. Dann gilt mit  $V = U(K) \cdots U(1)$ , dass

$$V(A|\mathbf{1}_N) = U(K) \cdots U(k+1)(O|B) = (\mathbf{1}_N|C),$$

Letzteres nach Definition von  $C$ . Nun gilt nach den Regeln der Matrizenmultiplikation  $V(A|\mathbf{1}_N) = (VA|V)$  und somit

$$(VA|V) = (\mathbf{1}_N|C).$$

Die erste Gleichung besagt also  $VA = \mathbf{1}_N$ , d.h.  $V = A^{-1}$ , die zweite, dass  $V = C$ , zusammen also tatsächlich  $C = A^{-1}$ .  $\square$

**Alternativer Beweis.** Gesucht ist eine Matrix  $X \in \text{Mat}(N \times N, \mathbb{K})$ , so dass  $AX = \mathbf{1}_N$ . Sei  $X$  aufgeteilt in seine Spaltenvektoren  $x_n \in \mathbb{K}^N$ , d.h.  $X = (x_1, \dots, x_N)$ . Auswerten auf der Standardbasis  $\{e_1, \dots, e_N\}$  ergibt  $Ax_n = e_n$ . Jede dieser  $N$  Gleichungen wird durch den Gauß-Algorithmus von  $(A|e_n)$  zu  $(\mathbf{1}_N|x_n)$  gelöst. Simultanes Lösen ist genau der beschriebene Algorithmus.  $\square$

**4.21 Korollar** Sei  $A \in \text{Mat}(N \times N, \mathbb{K})$  bijektiv. Dann ist die Lösung des linearen Gleichungssystems  $Ax = b$  mit  $b \in \mathbb{K}^N$  gegeben durch  $x = A^{-1}b$ .

**Beweis.** Hierzu wird lediglich die Gleichung  $Ax = b$  von links mit  $A^{-1}$  multipliziert.  $\square$

**4.22 Beispiel** Sei

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Die Schritte des Gauß-Algorithmus sind

$$\begin{array}{c|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array} \longrightarrow \begin{array}{c|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array} \longrightarrow \begin{array}{c|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array} .$$

Hieran kann abgelesen werden, dass der Rang gleich 2 ist, so dass tatsächlich invertiert werden kann. Der letzte Schritt ergibt:

$$\begin{array}{c|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array} .$$

Also ist

$$A^{-1} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix} .$$

Es ist sinnvoll die Probe zu machen:

$$AA^{-1} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} .$$

Es ist auch möglich obige Rechnung abstrakt durchzuführen, d.h. ohne Zahlen einzusetzen. Dies führt zu

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} ,$$

wobei  $a, b, c, d \in \mathbb{K}$  und vorausgesetzt wird, dass  $ad - bc \neq 0$  (sonst ist der Rang echt kleiner als 2).  $\diamond$

#### 4.23 Beispiel Sei

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix} .$$

Es soll das Inverse berechnet werden, falls es existiert. Mit dem Gauß-Algorithmus ergibt sich:

$$\begin{array}{c|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 2 & 1 & 0 & 0 & 0 & 1 \end{array} \longrightarrow \begin{array}{c|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -1 & 1 & 0 \\ 0 & -1 & -4 & -2 & 0 & 1 \end{array} \longrightarrow \begin{array}{c|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & -1 & 0 \\ 0 & 0 & -3 & -1 & -1 & 1 \end{array}$$

Hier erkennen wir, dass  $\text{rk}(A) = 3$  ist und die Matrix daher invertierbar ist. Weitere elementare Zeilenumformungen ergeben schließlich:

$$\begin{array}{c|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & -1 & 0 \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \end{array} \longrightarrow \begin{array}{c|ccc} 1 & 1 & 0 & \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 0 & \frac{2}{3} & -\frac{4}{3} & \frac{1}{3} \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \end{array} \longrightarrow \begin{array}{c|ccc} 1 & 0 & 0 & -\frac{1}{3} & \frac{2}{3} & \frac{1}{3} \\ 0 & 1 & 0 & \frac{2}{3} & -\frac{4}{3} & \frac{1}{3} \\ 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \end{array} ,$$

und somit

$$A^{-1} = \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{4}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \end{pmatrix} .$$

Die Probe zeigt:

$$AA^{-1} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix} \frac{1}{3} \begin{pmatrix} -1 & 2 & 1 \\ 2 & -4 & 1 \\ 1 & 1 & -1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \mathbf{1}_3 ,$$

so dass in der Tat das Inverse berechnet wurde.  $\diamond$



**4.24 Definition** Eine Matrix  $A \in \text{Mat}(N \times N, \mathbb{K})$  heißt obere Dreiecksmatrix genau dann, wenn  $A_{n,m} = 0$  für  $n > m$  und eine echte obere Dreiecksmatrix, falls  $A_{n,m} = 0$  für  $n \geq m$ . Analog sind untere Dreiecksmatrizen und echte untere Dreiecksmatrizen definiert.

**4.25 Satz** Folgende Eigenschaften gelten:

- (i) Linearkombinationen von oberen Dreiecksmatrizen sind obere Dreiecksmatrizen.
- (ii) Produkte von oberen Dreiecksmatrizen sind obere Dreiecksmatrizen.
- (iii) Inverse invertierbarer obere Dreiecksmatrizen sind obere Dreiecksmatrizen.
- (iv) Eine obere Dreiecksmatrix ist invertierbar genau dann, wenn all seine Diagonaleinträge  $A_{n,n}$  ungleich 0 sind.
- (v) Inverse und Produkte von oberen Dreiecksmatrizen mit positiver Diagonale sind obere Dreiecksmatrizen mit positiver Diagonale  $A_{n,n} > 0$ .

Analoge Aussagen gelten für untere Dreiecksmatrizen.

**Beweis.** Der Beweis sei als Übung durchgeführt. □

Aussage (iii) besagt, dass die invertierbaren oberen Dreiecksmatrizen eine Untergruppe der allgemeinen linearen Gruppe bilden. Nach (v) bilden die oberen Dreiecksmatrizen mit positiver Diagonale hiervon wieder eine Untergruppe. Nun beginnen die Vorbereitungen für die Koordinatenabbildungen und darstellenden Matrizen. Der folgende Sachverhalt hätte auch schon viel früher bewiesen werden können.

**4.26 Satz** Eine lineare Abbildung  $T \in \mathcal{L}(V, W)$  ist durch seine Werte auf einer Basis  $\mathcal{B} = \{b_1, b_2, \dots\}$  von  $V$  festgelegt, und zwar durch die Formel

$$T(v) = \sum_{n=1}^N v_n T(b_n), \quad \text{für } v = \sum_{n=1}^N v_n b_n.$$

**Beweis.** In die Formel gehen offensichtlich nur die Werte von  $T$  auf  $\mathcal{B}$  ein, und sie definiert auch eine Abbildung auf eindeutige Art und Weise, da jeder Vektor  $v$  gemäß Satz 2.18 auf eindeutige Art und Weise als endliche Linearkombination von Vektoren aus  $\mathcal{B}$  geschrieben werden kann. Wir überprüfen noch die Linearität. Sei  $w = \sum_{n=1}^N w_n b_n$  und  $\lambda \in \mathbb{K}$ . Dann gilt

$$T(v + \lambda w) = \sum_{n=1}^N (v_n + \lambda w_n) T(b_n), \quad \text{für } v + \lambda w = \sum_{n=1}^N (v_n + \lambda w_n) b_n,$$

und das Umsortieren der Summe ergibt:

$$T(v + \lambda w) = \sum_{n=1}^N v_n T(b_n) + \lambda \sum_{n=1}^N w_n T(b_n) = T(v) + \lambda T(w),$$

was den Beweis beendet. □

Der folgende Satz besagt, dass jeder endlich-dimensionale  $\mathbb{K}$ -Vektorraum letztendlich nur ein  $\mathbb{K}^N$  ist, zumindestens bis auf Isomorphie. Für unendlich dimensionale Vektorräume gibt es keinen solchen Satz, es sei denn man nimmt zusätzliche Strukturen hinzu (so ist jeder separable Hilbert-Raum zu einem Standardbeispiel isomorph). Es wird folgender Begriff benötigt.

**4.27 Definition** Eine Basis  $\mathcal{B} = \{b_1, b_2, \dots\}$  heißt geordnet, wenn eine Reihenfolge  $b_1, b_2, \dots$  festgelegt ist, d.h. eine Abbildung  $n \in \mathbb{N} \mapsto b_n \in \mathcal{B}$  oder  $n \in \{1, \dots, N\} \mapsto b_n \in \mathcal{B}$ . Für eine geordnete Basis werden die Bezeichnungen  $\mathcal{B} = (b_1, b_2, \dots) = (b_n)_{n=1,2,\dots}$  verwandt.

**4.28 Satz** Sei  $V$  ein Vektorraum über  $\mathbb{K}$  der Dimension  $N$ . Dann ist  $V \cong \mathbb{K}^N$  und zugehörig zu jeder geordneten Basis  $\mathcal{B} = (b_1, \dots, b_N)$  gibt es einen Isomorphismus  $I_{\mathcal{B}} : V \rightarrow \mathbb{K}^N$  definiert durch

$$I_{\mathcal{B}}(v) = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix}, \quad \text{für } v = \sum_{n=1}^N v_n b_n.$$

Dann wird  $I_{\mathcal{B}}(v)$  auch als der Koordinatenvektor von  $v$  bez.  $\mathcal{B}$  bezeichnet, und  $I_{\mathcal{B}}$  als die Koordinatenabbildung. Ihr Inverses  $(I_{\mathcal{B}})^{-1} : \mathbb{K}^N \rightarrow V$  ist gegeben durch

$$(I_{\mathcal{B}})^{-1} \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} = \sum_{n=1}^N v_n b_n.$$

**Beweis.** Die Definition und Linearität von  $I_{\mathcal{B}}$  ist ein Spezialfall von Satz 4.26, wobei der Wert des Basisvektors  $b_n$  gewählt wird als Standardbasisvektor  $e_n$ :

$$I_{\mathcal{B}}(b_n) = e_n.$$

In der Tat gilt dann

$$I_{\mathcal{B}}(v) = \sum_{n=1}^N v_n I_{\mathcal{B}}(b_n) = \sum_{n=1}^N v_n e_n = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix}, \quad \text{für } v = \sum_{n=1}^N v_n b_n.$$

Es verbleibt noch zu notieren, dass  $I_{\mathcal{B}}$  tatsächlich ein Isomorphismus ist. Gemäß Bemerkung 4.16 ist dies äquivalent dazu, dass eine Basis auf eine Basis abgebildet wird, was hier genau die Definition von  $I_{\mathcal{B}}$  ist.  $\square$

**4.29 Bemerkung** Die Gleichung  $I_{\mathcal{B}}(b_n) = e_n$  kann auch geschrieben werden als  $b_n = (I_{\mathcal{B}})^{-1}(e_n)$  und somit

$$(I_{\mathcal{B}})^{-1} = (b_1, \dots, b_N),$$

wobei auf der rechten Seite  $b_n$  als der  $n$ -te Spaltenvektor mit Werten in  $V$  ist. Falls  $V = \mathbb{K}^N$ , so ist  $(I_{\mathcal{B}})^{-1}$  eine  $N \times N$ -Matrix (d.h. dass hier  $(b_1, \dots, b_N)$  eine Matrix und nicht eine geordnete Basis bezeichnet).  $\diamond$

Nun können darstellende Matrizen eingeführt werden.

**4.30 Definition** Seien  $\mathcal{B} = (b_1, \dots, b_M)$  und  $\mathcal{C} = (c_1, \dots, c_N)$  geordnete Basen von zwei endlich-dimensionalen  $\mathbb{K}$ -Vektorräumen  $V$  und  $W$ . Zu  $T \in \mathcal{L}(V, W)$  wird nun die darstellende Matrix  $T_{\mathcal{C}}^{\mathcal{B}} \in \text{Mat}(N \times M, \mathbb{K})$  bez.  $\mathcal{B}$  und  $\mathcal{C}$  definiert durch

$$T_{\mathcal{C}}^{\mathcal{B}} = I_{\mathcal{C}} T (I_{\mathcal{B}})^{-1}. \quad (4.2)$$

**4.31 Bemerkung** Die in (4.2) definierte Abbildung  $T_{\mathcal{C}}^{\mathcal{B}}$  ist als Hintereinanderausführung von 3 linearen Abbildungen auch linear (nach Satz 4.13(i)). Außerdem sind Bild und Urbildraum gegeben durch  $T_{\mathcal{C}}^{\mathcal{B}} : \mathbb{K}^M \rightarrow \mathbb{K}^N$ . Da gemäß Satz 4.26 jede lineare Abbildung durch ihre Werte auf einer Basis festgelegt ist

und diese Basis als die Standardbasis im  $\mathbb{K}^M$  gewählt wird, werden also lediglich die Vektoren  $T_C^{\mathcal{B}}(e_m)$ ,  $m = 1, \dots, M$ , benötigt. Diese Vektoren bilden die Spaltenvektoren einer Matrix, eben der darstellenden Matrix  $T_C^{\mathcal{B}}$ . Die Einträge  $(T_C^{\mathcal{B}})_{n,m} \in \mathbb{K}$  der darstellenden Matrix

$$T_C^{\mathcal{B}} = \left( (T_C^{\mathcal{B}})_{n,m} \right)_{n=1, \dots, N, m=1, \dots, M},$$

sind also gegeben durch die Formel

$$T_C^{\mathcal{B}}(e_m) = \begin{pmatrix} (T_C^{\mathcal{B}})_{1,m} \\ \vdots \\ (T_C^{\mathcal{B}})_{N,m} \end{pmatrix}.$$

Da  $I_{\mathcal{B}}(b_m) = e_m$ , folgt durch Einsetzen von (4.2)

$$I_C T(b_m) = \begin{pmatrix} (T_C^{\mathcal{B}})_{1,m} \\ \vdots \\ (T_C^{\mathcal{B}})_{N,m} \end{pmatrix}.$$

Wenn nun von links  $(I_C)^{-1}$  angewandt wird, erhält man

$$T(b_m) = \sum_{n=1}^N (T_C^{\mathcal{B}})_{n,m} c_n, \tag{4.3}$$

was auch alternativ zur Definition der Einträge von  $T_C^{\mathcal{B}}$  verwandt werden kann. Nun kann die Matrix-

Vektormultiplikation von  $\begin{pmatrix} v_1 \\ \vdots \\ v_M \end{pmatrix} = I_{\mathcal{B}}(v)$  mit  $T_C^{\mathcal{B}}$  berechnet werden:

$$\begin{aligned} T_C^{\mathcal{B}} \begin{pmatrix} v_1 \\ \vdots \\ v_M \end{pmatrix} &= \begin{pmatrix} \sum_{m=1}^M (T_C^{\mathcal{B}})_{1,m} v_m \\ \vdots \\ \sum_{m=1}^M (T_C^{\mathcal{B}})_{N,m} v_m \end{pmatrix} = \sum_{m=1}^M v_m \begin{pmatrix} (T_C^{\mathcal{B}})_{1,m} \\ \vdots \\ (T_C^{\mathcal{B}})_{N,m} \end{pmatrix} \\ &= \sum_{m=1}^M v_m I_C T(b_m) = I_C T \left( \sum_{m=1}^M v_m b_m \right) = I_C T(v). \end{aligned}$$

Somit ist schon der folgende Satz bewiesen. ◇

**4.32 Satz** Seien  $V$  und  $W$  endlich dimensionale  $\mathbb{K}$ -Vektorräume mit geordneten Basen  $\mathcal{B}$  und  $\mathcal{C}$ . Dann gilt für  $T \in \mathcal{L}(V, W)$

$$I_C(Tv) = T_C^{\mathcal{B}} I_{\mathcal{B}}(v), \quad v \in V,$$

wobei auf der rechten Seite die Matrix-Vektor-Multiplikation verwandt wird.

**4.33 Bemerkung** Am einfachsten kann sich dieser Satz und die Definition der darstellenden Matrix mit Hilfe eines sogenannten kommutativen Diagrammes gemerkt werden:

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ I_{\mathcal{B}} \downarrow & & \downarrow I_{\mathcal{C}} \\ \mathbb{K}^M & \xrightarrow{T_C^{\mathcal{B}}} & \mathbb{K}^N \end{array}$$

wobei in der unteren Zeile die Matrixmultiplikation mit der Matrix  $T_C^{\mathcal{B}}$  gemeint ist. Die Kommutativität bedeutet dabei, dass von oben links nach unten rechts beiden Wegen gefolgt werden kann und dies zur gleichen Abbildung führt.  $\diamond$

**4.34 Beispiele** 1. Eine Matrix  $A \in \text{Mat}(N \times M, \mathbb{K})$  wurde ja mit einer linearen Abbildung  $T_A : \mathbb{K}^M \rightarrow \mathbb{K}^N$  identifiziert, die durch die Matrix-Vektor-Multiplikation definiert ist. Selbstverständlich ist die darstellende Matrix von  $T_A$  bez. der Standardbasen  $\mathcal{B} = (e_1, \dots, e_M)$  und  $\mathcal{C} = (e_1, \dots, e_N)$  gegeben durch die Matrix  $A$  selber:

$$(T_A)_{\mathcal{C}}^{\mathcal{B}} = A.$$

Dies im Detail zu überprüfen ist eine Übung.

2. Sei  $V = \mathbb{C}_N[x]$  der komplexe Vektorraum der komplexen Polynome in  $x$  vom Grad kleiner gleich  $N \in \mathbb{N}$ . Dieser Vektorraum hat die geordnete Basis  $\mathcal{B} = (b_0, \dots, b_N)$  mit  $b_n(x) = x^n$  und somit gibt es einen Isomorphismus  $I_{\mathcal{B}} : \mathbb{C}_N[x] \rightarrow \mathbb{C}^{N+1}$ . Betrachtet wird die Ableitung  $\partial_x : V \rightarrow V$ , deren Linearität schon in Beispiel 4.2 nachgewiesen wurde. Nun ist die darstellende Matrix  $A = (\partial_x)_{\mathcal{B}}^{\mathcal{B}} \in \text{Mat}((N+1) \times (N+1), \mathbb{C})$  gesucht. Hierzu wird die Identität

$$\partial_x b_n = n b_{n-1},$$

wobei insbesondere  $\partial_x b_0 = 0$ . Somit impliziert (4.3)

$$A = \begin{pmatrix} 0 & 1 & 0 & & & \\ & 0 & 2 & 0 & & \\ & & 0 & 3 & 0 & \\ & & & \ddots & \ddots & \\ & & & & 0 & N \\ & & & & & 0 \end{pmatrix},$$

oder anders ausgedrückt, die Koeffizienten von  $A$  sind  $A_{n,m} = (n+1) \delta_{n+1,m}$  für  $n, m = 0, \dots, N$ , wobei  $\delta$  das schon früher definierte Kronecker Delta ist. Wenn  $\partial_x$  als lineare Abbildung auf  $\mathbb{R}_N[x]$  aufgefasst wird, kann die gleiche Basis verwandt werden und die darstellende Matrix ist auch die gleiche.  $\diamond$

**4.35 Satz** Seien  $V$  und  $W$  Vektorräume über  $\mathbb{K}$  mit endlichen Dimensionen  $M = \dim(V)$  und  $N = \dim(W)$ , und  $\mathcal{B}$  und  $\mathcal{C}$  jeweils geordnete Basen. Dann:

(i)  $\dim(\mathcal{L}(V, W)) = \dim(V) \dim(W)$ .

(ii) Die Abbildung  $\mathcal{A} : \mathcal{L}(V, W) \rightarrow \text{Mat}(N \times M, \mathbb{K})$  definiert durch  $\mathcal{A}(T) = T_{\mathcal{C}}^{\mathcal{B}}$  ist ein Vektorraum Isomorphismus und somit

$$\mathcal{L}(V, W) \cong \text{Mat}(N \times M, \mathbb{K}).$$

(iii) Falls  $V = W$  und  $\mathcal{B} = \mathcal{C}$ , ist  $\mathcal{A}$  ein Ring-Homomorphismus. Insbesondere gilt

$$(ST)_{\mathcal{B}}^{\mathcal{B}} = S_{\mathcal{B}}^{\mathcal{B}} T_{\mathcal{B}}^{\mathcal{B}}, \quad S, T \in \mathcal{L}(V, W),$$

und für invertierbares  $T$

$$(T^{-1})_{\mathcal{B}}^{\mathcal{B}} = (T_{\mathcal{B}}^{\mathcal{B}})^{-1}.$$

(iv) Sei  $U$  ein weiterer  $\mathbb{K}$ -Vektorraum mit geordneter Basis  $\mathcal{D}$ , dann gilt

$$(ST)_{\mathcal{D}}^{\mathcal{B}} = S_{\mathcal{D}}^{\mathcal{C}} T_{\mathcal{C}}^{\mathcal{B}}, \quad T \in \mathcal{L}(V, W), \quad S \in \mathcal{L}(W, U),$$

d.h. "C kürzt sich raus".

**Beweis.** (i) Sei  $\mathcal{B} = (b_1, \dots, b_M)$  und  $\mathcal{C} = (c_1, \dots, c_N)$ . Für  $m = 1, \dots, M$  und  $n = 1, \dots, N$ , definieren wir (unter Verwendung von Satz 4.26)  $T^{(n,m)} \in \mathcal{L}(V, W)$  durch

$$T^{(n,m)}(b_j) = \delta_{j,m} c_n, \quad j = 1, \dots, M.$$

Diese  $NM$  linearen Abbildungen bilden eine Basis von  $\mathcal{L}(V, W)$  (Übung), so dass (i) folgt. (ii) Zunächst ist  $\mathcal{A}$  linear, da nach (4.2) und unter Verwendung der Linearität von  $I_{\mathcal{B}}$  und  $I_{\mathcal{C}}$

$$(T + \lambda S)_{\mathcal{C}}^{\mathcal{B}} = I_{\mathcal{C}}(T + \lambda S)(I_{\mathcal{B}})^{-1} = I_{\mathcal{C}}T(I_{\mathcal{B}})^{-1} + \lambda I_{\mathcal{C}}S(I_{\mathcal{B}})^{-1} = T_{\mathcal{C}}^{\mathcal{B}} + \lambda S_{\mathcal{C}}^{\mathcal{B}}, \quad T, S \in \mathcal{L}(V, W).$$

Außerdem ist  $T_{\mathcal{C}}^{\mathcal{B}} = 0$  genau dann, wenn  $T = 0$ , siehe z.B. (4.3). Somit ist der Kern von  $\mathcal{A}$  trivial und deswegen ist  $\mathcal{A}$  injektiv. Nach dem Rangsatz gilt nun  $\dim(\text{Ran}(\mathcal{A})) = \dim(\mathcal{L}(V, W)) = NM$ , Letzteres nach (i). Da aber auch  $\dim(\text{Mat}(N \times M, \mathbb{K})) = NM$ , folgt dass  $\mathcal{A}$  auch surjektiv ist. Also ist  $\mathcal{A}$  ein Isomorphismus. (iii) ist ein Spezialfall von (iv), welches durch mehrfache Anwendung von Satz 4.32 wie folgt nachgewiesen wird. Für  $v \in V$  gilt

$$I_{\mathcal{D}}(STv) = (ST)_{\mathcal{D}}^{\mathcal{B}} I_{\mathcal{B}}(v) = S_{\mathcal{D}}^{\mathcal{C}} I_{\mathcal{C}}(Tv) = S_{\mathcal{D}}^{\mathcal{C}} T_{\mathcal{C}}^{\mathcal{B}} I_{\mathcal{B}}(v),$$

was die Behauptung zeigt. Ein alternativer Beweis besteht im Aneinanderfügen von zwei kommutativen Diagrammen.  $\square$

Nun kommen wir zu Basiswechseln.

**4.36 Satz** Seien  $V$  und  $W$  Vektorräume über  $\mathbb{K}$  mit endlichen Dimensionen  $M = \dim(V)$  und  $N = \dim(W)$  und seien  $\mathcal{B}, \mathcal{B}'$  Basen von  $V$  und  $\mathcal{C}, \mathcal{C}'$  Basen von  $W$ . Für  $T \in \mathcal{L}(V, W)$  gilt dann

$$T_{\mathcal{C}'}^{\mathcal{B}'} = (\mathbf{1}_W)_{\mathcal{C}'}^{\mathcal{C}} T_{\mathcal{C}}^{\mathcal{B}} (\mathbf{1}_V)_{\mathcal{B}}^{\mathcal{B}'},$$

wobei  $\mathbf{1}_V$  und  $\mathbf{1}_W$  die Identitätsabbildungen auf  $V$  und  $W$  sind. Die  $M \times M$  darstellende Matrix  $(\mathbf{1}_V)_{\mathcal{B}}^{\mathcal{B}'}$  heißt Basiswechsel von  $\mathcal{B}'$  nach  $\mathcal{B}$ . Analog heißt die  $N \times N$  Matrix  $(\mathbf{1}_W)_{\mathcal{C}'}^{\mathcal{C}}$  Basiswechsel von  $\mathcal{C}$  nach  $\mathcal{C}'$ . Es gilt jeweils

$$((\mathbf{1}_V)_{\mathcal{B}}^{\mathcal{B}'})^{-1} = (\mathbf{1}_V)_{\mathcal{B}'}^{\mathcal{B}} \quad ((\mathbf{1}_W)_{\mathcal{C}'}^{\mathcal{C}})^{-1} = (\mathbf{1}_W)_{\mathcal{C}}^{\mathcal{C}'}$$

**Beweis.** Die erste Aussage folgt direkt aus iterativer Anwendung von Satz 4.35(iv), die zweite Aussage dann aus Satz 4.35(iii)  $\square$

**4.37 Bemerkung** Gemäß der Definition darstellender Matrizen gilt

$$(\mathbf{1}_V)_{\mathcal{B}}^{\mathcal{B}'} = I_{\mathcal{B}}(I_{\mathcal{B}'})^{-1}.$$

Wenn nun  $V = \mathbb{K}^N$  und  $\mathcal{B} = (e_1, \dots, e_N)$  die Standardbasis ist, so ist  $I_{\mathcal{B}} = \mathbf{1}_N$  die  $N \times N$  Einheitsmatrix. Wenn des Weiteren  $\mathcal{B}' = (b_1, \dots, b_N)$  eine weitere Basis des  $\mathbb{K}^N$  ist, so gilt nach Bemerkung 4.37, dass  $(I_{\mathcal{B}'})^{-1} = (b_1, \dots, b_N)$  und somit

$$(\mathbf{1}_V)_{\mathcal{B}}^{\mathcal{B}'} = (b_1, \dots, b_N).$$

Diese Matrix ist invertierbar. Umgekehrt können die Spaltenvektoren einer invertierbaren Matrix als Elemente einer Basis aufgesest werden. Somit gibt es eine Bijektion zwischen der Menge  $\{\mathcal{B}' \text{ Basis von } \mathbb{K}^N\}$  der Basen von  $\mathbb{K}^N$  und der allgemeinen linearen Gruppe  $\text{Gl}(N, \mathbb{K})$  der invertierbaren  $N \times N$  Matrizen mit Einträgen in  $\mathbb{K}$ .  $\diamond$

**4.38 Beispiel** Im  $V = \mathbb{R}^3$  sei  $\mathcal{B}$  die geordnete Standardbasis, und eine zweite geordnete Basis  $\mathcal{B}'$  gegeben durch

$$\mathcal{B}' = \left( \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right).$$

Dann ist

$$(\mathbf{1}_V)_{\mathcal{B}'}^{\mathcal{B}} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix}.$$

Des weiteren sei  $W = \mathbb{R}^2$  und  $\mathcal{C} = (e_1, e_2)$  die Standardbasis und

$$\mathcal{C}' = \left( \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right).$$

Dann ist

$$(\mathbf{1}_W)_{\mathcal{C}'}^{\mathcal{C}} = ((\mathbf{1}_W)_{\mathcal{C}}^{\mathcal{C}'})^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Nun sei gegeben die lineare Abbildung  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  durch

$$T \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} v_1 + v_3 \\ 2v_2 - v_3 \end{pmatrix}.$$

Bezüglich der Standardbasen  $\mathcal{B}$  und  $\mathcal{C}$  ist also die darstellende Matrix

$$T_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & -1 \end{pmatrix}.$$

Die Abbildung in den neuen Basen  $\mathcal{B}'$  und  $\mathcal{C}'$  ist dann gegeben durch

$$T_{\mathcal{C}'}^{\mathcal{B}'} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Selbstverständlich ist auch eine direkte Berechnung der darstellenden Matrix  $T_{\mathcal{C}'}^{\mathcal{B}'}$  ausgehend von der Definition möglich.  $\diamond$

Ein Spezialfall von Satz 4.36 betrifft eine Abbildung  $T \in \mathcal{L}(V) = \mathcal{L}(V, V)$  auf  $V$  und zwei Basen  $\mathcal{B}$  und  $\mathcal{B}'$  von  $V$ :

$$T_{\mathcal{B}'}^{\mathcal{B}'} = \mathbf{1}_{\mathcal{B}'}^{\mathcal{B}} T_{\mathcal{B}}^{\mathcal{B}} (\mathbf{1}_{\mathcal{B}'}^{\mathcal{B}})^{-1}.$$

Diese Gleichung besagt, dass die beiden darstellenden Matrizen  $T_{\mathcal{B}'}^{\mathcal{B}'}$  und  $T_{\mathcal{B}}^{\mathcal{B}}$  ähnlich sind im Sinne folgender Definition.

**4.39 Definition** Zwei Matrizen  $A, B \in \text{Mat}(N \times N, \mathbb{K})$  heißen *ähnlich genau dann*, wenn es ein invertierbare Matrix  $M \in \text{Mat}(N \times N, \mathbb{K})$  gibt mit

$$A = MBM^{-1}. \tag{4.4}$$

**4.40 Satz** Ähnlichkeit ist eine Äquivalenzrelation auf  $\text{Mat}(N \times N, \mathbb{K})$ .

**Beweis.** Die Ähnlichkeit sei mit  $\sim$  bezeichnet. Offensichtlich ist  $A \sim A$  (Reflexivität von  $\sim$ ), denn (4.4) gilt mit  $M = \mathbf{1}_N$ . Außerdem gilt  $A = MBM^{-1}$  genau dann, wenn  $M^{-1}AM = B$ , was gleichbedeutend mit der Symmetrie von  $\sim$  ist. Nun sei  $A \sim B$  und  $B \sim C$ . Dann gilt  $A = MBM^{-1}$  und  $B = NCN^{-1}$  für zwei invertierbare Matrizen  $N, M$ . Einsetzen ergibt:

$$A = MNCN^{-1}M^{-1} = (MN)C(MN)^{-1},$$

so dass auch  $A \sim C$ , d.h.  $\sim$  ist transitiv. □

Ein ganz wesentliches Ziel der linearen Algebra ist es innerhalb einer Äquivalenzklasse einen besonderen einfachen Repräsentanten zu finden, indem ein adäquater Basiswechsel  $M$  bestimmt wird. Dieser Repräsentant wird genau die Jordan'sche Normalform der Matrix sein.

## 5 Permutationen und die Determinante

**5.1 Definition** Sei  $N \in \mathbb{N}$ . Eine bijektive Abbildung  $\sigma : \{1, \dots, N\} \rightarrow \{1, \dots, N\}$  heißt eine Permutation über  $N$  Elementen. Die Menge  $S_N$  aller Permutationen über  $N$  Elementen versehen mit der Hintereinanderausführung heißt die symmetrische Gruppe. Eine Standardnotation für eine Permutation ist

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & N \\ \sigma(1) & \sigma(2) & \dots & \sigma(N) \end{pmatrix},$$

wobei die untere Zeile die der oberen Zeile zugeordneten Abbildungswerte enthält. Eine Transposition (von  $n', m'$ ) ist eine Permutation, bei der  $\sigma(n) \neq n$  für lediglich zwei Werte  $n = n', m'$ , bei denen  $\sigma(n') = m'$ . Bei einer benachbarten Transposition gilt zudem  $m' = n' + 1 \pmod N$ .

**5.2 Beispiel** Die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}$$

ist eine benachbarte Transposition. ◇

**5.3 Satz** (i) Die Gruppe  $S_N$  hat  $N! = N(N-1) \cdots 1$  Elemente.

(ii) Jede Transposition ist Hintereinanderausführung einer ungeraden Anzahl von benachbarten Transpositionen.

(iii) Jede Permutation ist Hintereinanderausführung von benachbarten Transpositionen.

**Beweis.** (i) Bei der Konstruktion einer Bijektion kann man zunächst  $\sigma(1)$  genau  $N$  Werte annehmen, danach kann  $\sigma(2)$  nur noch  $N-1$  Werte annehmen, etc. (ii) Seien  $n' < m'$  die durch die Transposition vertauschten Elemente. Zunächst werden  $m' - n'$  benachbarte Transpositionen ausgeführt, um  $n'$  an die Stelle  $m'$  zu bewegen, dann  $m' - n' - 1$  benachbarte Transpositionen, um  $m'$  an die Stelle  $n'$  zu bewegen. (iii) Nach (ii) reicht es jede Permutation als Hintereinanderausführung von Transpositionen zu schreiben. Dies wird durch Induktion über  $N$  gezeigt. Für  $N = 1, 2$  ist es offensichtlich richtig. Sei nun  $\sigma \in S_{N+1}$ . Entweder  $\sigma$  ist die Identität und nichts ist zu zeigen, oder es gibt ein  $n$  mit  $\sigma(n) = m \neq n$ . Sei  $\tau$  die Transposition von  $n$  und  $m$ . Dann gilt  $\sigma\tau(n) = n$ . Also ist  $\sigma\tau$  eingeschränkt auf  $\{1, \dots, N+1\} \setminus \{n\}$  eine Permutation von  $N$  Elementen, die nach Induktionsvoraussetzung also als Produkt von Hintereinanderausführungen geschrieben werden kann. □

**5.4 Satz** Zu jedem  $\sigma \in S_N$  sei seine darstellende Matrix  $P(\sigma) \in \text{Mat}(N \times N, \mathbb{R})$  definiert durch

$$P(\sigma)_{n,m} = \delta_{n,\sigma(m)}.$$

Mit Hilfe der Standardbasisvektoren als Spaltenvektoren, kann diese Matrix auch geschrieben werden als

$$P(\sigma) = (e_{\sigma(1)}, \dots, e_{\sigma(N)}).$$

Es gilt dann für  $\sigma, \tau \in S_N$

$$P(\sigma\tau) = P(\sigma)P(\tau),$$

wobei auf der rechten Seite die Matrixmultiplikation verwendet wird. Somit ist  $P(S_N)$  eine Untergruppe der allgemeinen linearen Gruppe  $\text{Gl}(N, \mathbb{R})$ .

**Beweis.** Es gilt nach Definition des Matrixproduktes

$$(P(\sigma)P(\tau))_{n,m} = \sum_{k=1}^N P(\sigma)_{n,k}P(\tau)_{k,m} = \sum_{k=1}^N \delta_{n,\sigma(k)}\delta_{k,\tau(m)} = \delta_{n,\sigma(\tau(m))},$$

was die erste Behauptung zeigt. Die zweite folgt dann aus Satz 1.8(ii). □

**5.5 Definition** Sei  $\sigma \in S_N$ . Ein Fehlstand von  $\sigma$  ist dann ein Paar  $(n, m)$  mit  $1 \leq n < m \leq N$ , so dass  $\sigma(n) > \sigma(m)$ . Die Anzahl  $L(\sigma)$  der Fehlstände von  $\sigma$  heißt auch die Länge von  $\sigma$ . Das Signum von  $\sigma$  ist definiert als

$$\text{sgn}(\sigma) = (-1)^{L(\sigma)}.$$

**5.6 Beispiel** Die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}$$

hat Fehlstände  $(1, 3)$ ,  $(2, 3)$  und  $(2, 4)$ . Also ist  $L(\sigma) = 3$  und  $\text{sgn}(\sigma) = -1$ . ◇

**5.7 Bemerkung** Für jede benachbarte Transposition  $\tau \in S_N$  gilt  $\text{sgn}(\tau) = -1$ , weil sie nämlich genau einen Fehlstand hat. ◇

**5.8 Satz** (i) Für  $\sigma \in S_N$ ,

$$\text{sgn}(\sigma) = \prod_{1 \leq n < m \leq N} \frac{\sigma(n) - \sigma(m)}{n - m},$$

wobei das Produkt gebildet wird über alle Paare  $(n, m)$  mit der angegebenen Eigenschaft  $1 \leq n < m \leq N$ .

(ii) Für  $\sigma, \tau \in S_N$ ,

$$\text{sgn}(\sigma\tau) = \text{sgn}(\sigma)\text{sgn}(\tau).$$

(iii) Die Abbildung  $\sigma \in S_N \mapsto \text{sgn}(\sigma) \in \mathbb{Z}_2$  ist ein surjektiver Gruppen-Homomorphismus, wobei hier  $\mathbb{Z}_2$  die multiplikative Gruppe mit 2 Elementen ist.

(iv) Sei  $\sigma = \tau_K \cdots \tau_1$  ein Produkt von Transpositionen (was nach Satz 5.3(iii) immer der Fall ist), so ist  $\text{sgn}(\sigma) = (-1)^K$ .



**Beweis.** (i) Da  $\sigma$  eine Permutation ist, gibt es für jedes Paar  $(n, m)$  mit  $n < m$  genau ein (nicht geordnetes) Paar  $n', m'$  mit  $\sigma(n') = n$  und  $\sigma(m') = m$ . Also sind die Beträge der beiden Ausdrücke

$$\prod_{1 \leq n < m \leq N} (n - m), \quad \prod_{1 \leq n' < m' \leq N} (\sigma(n') - \sigma(m')),$$

gleich. Des Weiteren gilt  $n' < m'$ , falls kein Fehlstand von  $\sigma$  bei  $(n', m')$  vorliegt, und  $m' < n'$ , falls einer vorliegt. In ersterem Fall sind  $n - m$  und  $\sigma(n') - \sigma(m')$  beide negativ, wohingegen in zweiterem  $n - m$  negativ und  $\sigma(n') - \sigma(m')$  positiv ist. Somit gilt

$$\prod_{1 \leq n < m \leq N} (n - m) = (-1)^{L(\sigma)} \prod_{1 \leq n' < m' \leq N} (\sigma(n') - \sigma(m')).$$

Die Variablen  $n', m'$  auf der rechten Seite können durch  $n, m$  ersetzt werden, und die beiden Produkte zusammengefasst werden. Das zeigt dann die Formel. (ii) Unter Verwendung der Formel folgt nun

$$\begin{aligned} \operatorname{sgn}(\sigma\tau) &= \prod_{1 \leq n < m \leq N} \frac{\sigma(\tau(n)) - \sigma(\tau(m))}{n - m} \\ &= \prod_{1 \leq n < m \leq N} \frac{\sigma(\tau(n)) - \sigma(\tau(m))}{\tau(n) - \tau(m)} \prod_{1 \leq n < m \leq N} \frac{\tau(n) - \tau(m)}{n - m} \\ &= \operatorname{sgn}(\sigma) \operatorname{sgn}(\tau). \end{aligned}$$

(iii) ist genau der in (ii) beschriebene Sachverhalt und (iv) folgt auch aus (ii), da jede Transposition ein negatives Signum hat, da sie ja nach Satz 5.3(iii) aus einer ungeraden Anzahl von benachbarten Transpositionen besteht.  $\square$

Nach diesen Vorbereitungen kann nun die zentrale Definition dieses Kapitels folgen:

**5.9 Definition** Sei  $A = (A_{n,m})_{n,m=1,\dots,N} \in \operatorname{Mat}(N \times N, \mathbb{K})$ . Seine Determinante ist dann definiert als

$$\det(A) = \sum_{\sigma \in S_N} \operatorname{sgn}(\sigma) \prod_{n=1}^N A_{n,\sigma(n)} = \sum_{\sigma \in S_N} \operatorname{sgn}(\sigma) A_{1,\sigma(1)} A_{2,\sigma(2)} \cdots A_{N,\sigma(N)}.$$

**5.10 Beispiel** Sei  $N = 2$ . Dann gibt es zwei Permutationen, die Identität und die Transposition von 1 mit 2. Somit gilt

$$\det(A) = A_{1,1}A_{2,2} - A_{1,2}A_{2,1}.$$

Wenn  $\mathbb{K} = \mathbb{R}$ ,  $a = \begin{pmatrix} A_{1,1} \\ A_{2,1} \end{pmatrix}$  und  $b = \begin{pmatrix} A_{1,2} \\ A_{2,2} \end{pmatrix}$ , so dass  $A = (a, b)$ , dann ist die Fläche des von  $a$  und  $b$  aufgespannten Parallelogramms gegeben durch den Betrag der Determinante. In der Tat, wenn  $\alpha$  der Winkel zwischen  $a$  und  $b$  ist, so ist die Fläche  $\sin(\alpha)\|a\|\|b\|$ , wobei  $\|a\| = ((A_{1,1})^2 + (A_{2,1})^2)^{\frac{1}{2}}$  die Länge von  $a$  ist, und entsprechend  $\|b\|$  die Länge von  $b$ . Dies kann mit aus der Schule bekannter Geometrie von Dreiecken überprüft werden, und im Rahmen der Vorlesung etwas später unter Verwendung von Skalarprodukten. Für  $3 \times 3$  Matrizen gilt die sogenannte Regel von Sarrus für die Berechnung von Determinanten, die in einer Übung diskutiert wird. Außerdem gilt ein entsprechender Sachverhalt über das Volumen des von den Spalten der Matrix aufgespannten Parallelipeds.  $\diamond$

**5.11 Beispiel** Sei  $\sigma \in S_N$  und  $P(\sigma)$  die zugehörige darstellende Matrix. Dann gilt

$$\det(P(\sigma)) = \operatorname{sgn}(\sigma),$$

da lediglich ein Summand in der Definition nicht verschwindet.  $\diamond$

**5.12 Bemerkung** Es ist auch möglich die Formel

$$\det(A) = \sum_{\sigma \in S_N} \operatorname{sgn}(\sigma) \prod_{n=1}^N A_{\sigma(n),n}$$

zu verwenden, denn ausgehend von der obigen Definition und der Bijektivität jeder Permutation gilt:

$$\det(A) = \sum_{\sigma \in S_N} \operatorname{sgn}(\sigma) \prod_{n=1}^N A_{\sigma^{-1}(n),n} .$$

Aber  $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$ , so dass nach Umbenennung obige Formel folgt. Die Aussage kann auch formuliert werden als

$$\det(A^t) = \det(A) ,$$

wobei  $A^t$  die sogenannte transponierte Matrix im Sinne der folgenden Definition bezeichnet.  $\diamond$

**5.13 Definition** Sei  $A = (A_{n,m})_{n,1,\dots,N,m=1,\dots,M} \in \operatorname{Mat}(N \times M, \mathbb{K})$ . Seine transponierte Matrix  $A^t \in \operatorname{Mat}(M \times N, \mathbb{K})$  ist dann definiert durch

$$(A^t)_{m,n} = A_{n,m} .$$

Die wesentlichen algebraischen Eigenschaften der Determinanten sind ihre Multilinearität und Alternierendeneigenschaft, welche nun etwas allgemeiner definiert werden. Zusammen mit einer geeigneten Normierung charakterisieren diese Eigenschaften sogar die Determinante, wie wir später zeigen werden.

**5.14 Definition** Seien  $V_1, \dots, V_N$  und  $W$  Vektorräume über  $\mathbb{K}$ . Eine Abbildung  $F : V_1 \times \dots \times V_N \rightarrow W$  heißt *multilinear genau dann*, wenn sie *linear in jedem Argument ist, d.h. wenn für jedes  $n = 1, \dots, N$  die Abbildung  $v_n \in V_n \mapsto F(v_1, \dots, v_N) \in W$  linear ist für alle festgehaltenen  $v_m \in V_m, m \neq n$ . Falls  $N = 2$  und  $V_1 = V_2$ , wird auch von einer bilinearen Abbildung gesprochen.*

**5.15 Beispiel** Jede lineare Abbildung ist auch multilinear mit  $N = 1$ . Nun zu einem nicht-trivialen Beispiel. Sei  $V = \mathbb{K}^M$  und  $A \in \operatorname{Mat}(M \times M, \mathbb{K})$ . Dann ist  $F : V \times V \rightarrow \mathbb{K}$  definiert durch

$$F(v, w) = \sum_{m,n=1}^M v_m A_{m,n} w_n ,$$

eine bilineare Abbildung. Im Fall  $M = 3$  ist das Kreuzprodukt ein konkretes Beispiel dieser Natur, siehe Übung.  $\diamond$

**5.16 Definition** Seien  $V$  und  $W$  Vektorräume über  $\mathbb{K}$ . Eine Abbildung  $F : V^{\times N} \rightarrow W$  heißt *alternierend genau dann*, wenn für jede benachbarte Transposition  $\tau \in S_N$  (d.h. jedes Vertauschen von zwei benachbarten Argumenten) gilt

$$F(v_{\tau(1)}, \dots, v_{\tau(N)}) = -F(v_1, \dots, v_N) , \quad v_n \in V .$$

**5.17 Beispiel** Sei  $A \in \operatorname{Mat}(M \times M, \mathbb{K})$  eine anti-symmetrische Matrix, d.h.

$$A^t = -A .$$

Dann ist  $F : \mathbb{K}^M \times \mathbb{K}^M \rightarrow \mathbb{K}$  wie oben definiert durch

$$F(v, w) = \sum_{m,n=1}^M v_m A_{m,n} w_n ,$$

eine alternierende bilineare Abbildung. In der Tat,

$$F(w, v) = \sum_{m,n=1}^M w_m A_{m,n} v_n = - \sum_{m,n=1}^M w_m A_{n,m} v_n = -F(v, w) .$$

Wiederum ist im Fall  $M = 3$  das Kreuzprodukt ein konkretes Beispiel. ◇

**5.18 Lemma** Seien  $V$  und  $W$  Vektorräume über  $\mathbb{K}$  und  $F : V^{\times N} \rightarrow W$  alternierend. Dann gilt für jede Permutation  $\sigma \in S_N$

$$F(v_{\sigma(1)}, \dots, v_{\sigma(N)}) = \text{sgn}(\sigma) F(v_1, \dots, v_N) ,$$

für alle  $v_1, \dots, v_N \in V$ .

**Beweis.** Da gemäß Satz 5.3(ii) jede Permutation Hintereinanderausführung von benachbarten Transpositionen ist und die Anzahl der Transpositionen genau dem Signum entspricht, siehe Satz 5.8(iv), folgt die Aussage aus der Definition. □

**5.19 Satz** Sei  $F : V^{\times N} \rightarrow W$  eine multilineare, alternierende Abbildung.

(i) Falls dann  $v_1, \dots, v_N \in V$  linear abhängig sind, so gilt

$$F(v_1, \dots, v_N) = 0 .$$

Insbesondere ist  $F(v_1, \dots, v_N) = 0$ , falls  $v_n = v_m$  für  $n \neq m$ .

(ii) Sei  $\dim(V) = N$  und  $F \neq 0$ , d.h. es gibt  $w_1, \dots, w_N \in V$  mit  $F(w_1, \dots, w_N) \neq 0$ . Falls dann  $v_1, \dots, v_N \in V$  linear unabhängig sind, so gilt

$$F(v_1, \dots, v_N) \neq 0 .$$

**Beweis.** (i) Zunächst wird die zweite Behauptung gezeigt. Hierzu wird obiges Lemma auf die Transposition  $\tau$  von  $n$  und  $m$  angewandt, so dass  $F(\dots, v_m, \dots, v_n, \dots) = -F(\dots, v_n, \dots, v_m, \dots) = 0$ , falls  $v_n = v_m$ . Nun seien  $v_1, \dots, v_N$  linear abhängig, und z.B.  $v_N = \sum_{n=1}^{N-1} \lambda_n v_n$ . Dann folgt aus der Linearität im letzten Argument

$$F(v_1, \dots, v_N) = \sum_{n=1}^{N-1} \lambda_n F(v_1, \dots, v_{N-1}, v_n) = 0 ,$$

wobei Letzteres gilt, weil jeder Term in der Summe verschwindet. (ii) Nach den Voraussetzungen bilden  $v_1, \dots, v_N$  eine Basis von  $V$ . Somit können  $w_1, \dots, w_N$  in dieser Basis linear zerlegt werden:

$$w_n = \sum_{m=1}^N \lambda_{n,m} v_m .$$

Einsetzen liefert nun unter Verwendung der Multilinearität

$$F(w_1, \dots, w_N) = \sum_{m_1, \dots, m_N=1}^N (\lambda_{1,m_1} \cdots \lambda_{N,m_N}) F(v_{m_1}, \dots, v_{m_N}) .$$

Nach (i) verschwinden nun aber alle Summanden, für die  $m_i = m_j$  für ein Paar  $i \neq j$  gilt. Die verbleibenden Summanden erfüllen  $\{m_1, \dots, m_N\} = \{1, \dots, N\}$  und können somit durch die Permutation  $\sigma$  beschrieben werden, für die  $\sigma(n) = m_n$  gilt. Dann

$$\begin{aligned} F(w_1, \dots, w_N) &= \sum_{\sigma \in S_N} (\lambda_{1,\sigma(1)} \cdots \lambda_{N,\sigma(N)}) F(v_{\sigma(1)}, \dots, v_{\sigma(N)}) \\ &= \left( \sum_{\sigma \in S_N} (\lambda_{1,\sigma(1)} \cdots \lambda_{N,\sigma(N)}) \operatorname{sgn}(\sigma) \right) F(v_1, \dots, v_N), \end{aligned}$$

Letzteres nach Lemma 5.18. Da  $F(w_1, \dots, w_N) \neq 0$ , muss auch  $F(v_1, \dots, v_N) \neq 0$  gelten.  $\square$

**5.20 Satz** Die Abbildung  $A \in \operatorname{Mat}(N \times N, \mathbb{K}) \cong (\mathbb{K}^N)^{\times N} \mapsto \det(A) \in \mathbb{K}$  ist multilinear und alternierend, wenn aufgefasst als Abbildung auf den  $N$  Spaltenvektoren. Außerdem gilt  $\det(\mathbf{1}_N) = 1$ .

**Beweis.** Sei  $A = (a_1, \dots, a_N)$  mit Spaltenvektoren  $a_n \in \mathbb{K}^N$  für  $n = 1, \dots, N$ . Für  $\lambda \in \mathbb{K}$  und  $b \in \mathbb{K}^N$  gilt dann

$$\begin{aligned} \det(a_1, \dots, a_n + \lambda b, \dots, a_N) &= \sum_{\sigma \in S_N} \operatorname{sgn}(\sigma) A_{1,\sigma(1)} \cdots (A_{n,\sigma(n)} + \lambda b_{\sigma(n)}) \cdots A_{N,\sigma(N)} \\ &= \det(a_1, \dots, a_N) + \lambda \det(a_1, \dots, a_{n-1}, b, a_{n+1}, \dots, a_N), \end{aligned}$$

was schon die Linearität in jedem Argument ist. Des Weiteren sei  $\tau$  eine benachbarte Transposition. Dann gilt

$$\begin{aligned} \det(a_{\tau(1)}, \dots, a_{\tau(N)}) &= \sum_{\sigma \in S_N} \operatorname{sgn}(\sigma) A_{1,\sigma(\tau(1))} \cdots A_{N,\sigma(\tau(N))} \\ &= - \sum_{\sigma\tau \in S_N} \operatorname{sgn}(\sigma\tau) A_{1,\sigma(\tau(1))} \cdots A_{N,\sigma(\tau(N))}, \end{aligned}$$

wobei in der zweiten Gleichung die Homomorphismeigenschaft vom Signum verwandt wurde und  $\tau S_N = S_N$ . Umbenennung zeigt nun die Alternierendeigenschaft. Die letzte Behauptung folgt direkt durch Einsetzen (und ist ein Spezialfall von Beispiel 5.11).  $\square$

**5.21 Satz** Sei  $A \in \operatorname{Mat}(N \times N, \mathbb{K})$ . Dann gilt

$$A \text{ invertierbar} \iff \det(A) \neq 0.$$

**Beweis.** Nach Satz 5.20 ist die Determinante multilinear, alternierend und nicht identisch gleich Null, so dass die Aussagen aus Satz 5.19 gelten. “ $\implies$ ” Wenn  $A$  invertierbar ist, sind seine Zeilen linear unabhängig (denn der Zeilenrang ist gleich  $N$ ) und somit ist  $\det(A) \neq 0$  nach Satz 5.19(ii). “ $\impliedby$ ” Da  $A$  nicht invertierbar ist, so sind seine Zeilen linear abhängig und deswegen  $\det(A) = 0$  nach Satz 5.19(i).  $\square$

Nun folgt die Charakterisierung der Determinante, die oft auch zur Definition verwandt wird.

**5.22 Satz** Sei  $F : \operatorname{Mat}(N \times N, \mathbb{K}) \cong (\mathbb{K}^N)^{\times N} \rightarrow \mathbb{K}$ . Dann gilt:

$$F \text{ multilineare, alternierend mit } F(\mathbf{1}_N) = 1 \iff F = \det.$$

**Beweis.** "←" ist genau Satz 5.20. "⇒" Zunächst ist  $F$  wegen der Multilinearität schon durch seine Werte auf der Standardbasis  $e_1, \dots, e_N$  (in jedem Argument) festgelegt. Ein formaler Beweis hiervon sieht genau wie der von Satz 4.26 aus. Nun sind gemäß Satz 5.19(i) die Werte  $F(e_{i_1}, \dots, e_{i_N})$  nur ungleich Null, wenn  $\{i_1, \dots, i_N\} = \{1, \dots, N\}$ . Also gilt  $i_n = \sigma(n)$  für geeignetes  $\sigma \in S_N$ . Nach Lemma 5.18 gilt dann aber  $F(e_{i_1}, \dots, e_{i_N}) = \text{sgn}(\sigma)F(e_1, \dots, e_N)$ . Somit wird lediglich der Wert  $F(e_1, \dots, e_N)$  benötigt, der aber nach Voraussetzung genau  $1 = \det(\mathbf{1}_N)$  ist.  $\square$

Es werden nun eine Reihe von Eigenschaften der Determinante gezeigt.

**5.23 Satz** Seien  $A, B \in \text{Mat}(N \times N, \mathbb{K})$ . Dann gilt

$$\det(AB) = \det(A) \det(B).$$

Insbesondere gilt für invertierbares  $A$

$$\det(A^{-1}) = \frac{1}{\det(A)},$$

und

$$\det(A^{-1}BA) = \det(B).$$

**Beweis.** Sei  $A = (a_1, \dots, a_N)$  mit Spaltenvektoren  $a_n \in \mathbb{K}$ . Dann sind die Spaltenvektoren von  $AB$  gegeben durch

$$AB = \left( \sum_{n_1=1}^N B_{n_1,1} a_{n_1}, \dots, \sum_{n_1=1}^N B_{n_1,N} a_{n_1} \right).$$

Unter Verwendung der Multilinearität von  $\det$  folgt also

$$\begin{aligned} \det(AB) &= \det \left( \sum_{n_1=1}^N B_{n_1,1} a_{n_1}, \dots, \sum_{n_1=1}^N B_{n_1,N} a_{n_1} \right) \\ &= \sum_{n_1, \dots, n_N=1}^N B_{n_1,1} \cdots B_{n_N,N} \det(a_{n_1}, \dots, a_{n_N}). \end{aligned}$$

Hiervon verbleiben nur die Summanden mit  $\{n_1, \dots, n_N\} = \{1, \dots, N\}$ , für die es eine Permutation  $\sigma \in S_N$  gibt mit  $n_m = \sigma(m)$ . Somit

$$\det(AB) = \sum_{\sigma \in S_N} B_{\sigma(1),1} \cdots B_{\sigma(N),N} \det(a_{\sigma(1)}, \dots, a_{\sigma(N)}).$$

Nach Lemma 5.18 folgt nun

$$\det(AB) = \sum_{\sigma \in S_N} B_{\sigma(1),1} \cdots B_{\sigma(N),N} \text{sgn}(\sigma) \det(a_1, \dots, a_N) = \det(B^t) \det(A),$$

was wegen  $\det(B^t) = \det(B)$  den Beweis der Identität beendet. Die Zusätze folgen direkt daraus.  $\square$

**5.24 Bemerkung** Die Gleichung  $\det(A^{-1}BA) = \det(B)$  besagt, dass die Determinante von den ähnlichen Matrizen  $B$  und  $A^{-1}BA$  (vgl. Definition 4.39) gleich sind. Dies erlaubt auch die Determinante einer beliebigen linearen Abbildung  $T \in \mathcal{L}(V)$  auf einem endlich dimensionalen Vektorraum  $V$  zu definieren durch

$$\det(T) = \det(T_{\mathcal{B}}^{\mathcal{B}}),$$

wobei  $\mathcal{B}$  eine Basis von  $V$  ist und  $T_{\mathcal{B}}^{\mathcal{B}}$  die zugehörige darstellende Matrix. Obiges impliziert dann in der Tat, dass diese Definition unabhängig von der Wahl von  $\mathcal{B}$  ist.  $\diamond$

**5.25 Bemerkung** Eine Art der Berechnung der Determinante sei kurz diskutiert (Details in Übungen). Der Gauss-Algorithmus wird im Beweis von Satz 4.20 (Berechnung der inversen Matrix) beschrieben durch Linksmultiplikation mit drei Matrizen, die das Normieren, Vertauschen und Addieren anderer Zeilen implementieren. Die Determinanten dieser Matrizen können einfach abgelesen werden. Das Vertauschen führt lediglich zu einem Vorzeichen, wohingegen das Addieren des Vielfachen einer anderen Zeile Determinante 1 hat. Wenn also nicht normiert wird, so ändert sich die Determinante lediglich um einen Vorzeichenwechsel, was aus der Anzahl der benötigten Vertauschungen auch leicht bestimmt werden kann. Am Ende der ersten Etappe des Gauss-Algorithmus erhält man eine obere Dreiecksmatrix, deren Determinante gegeben ist durch das Produkt der Diagonaleinträge. Dies ist dann bis auf das Vorzeichen gleich der Determinante der ursprünglichen Matrix. Somit liefert das Gauss-Verfahren eine Methode, um die Determinante zu berechnen.  $\diamond$

**5.26 Beispiel** Mit dem Gauss-Algorithmus erhält man:

$$\det \left( \begin{pmatrix} 2 & 1 & -1 \\ 3 & 0 & 1 \\ 1 & 2 & 1 \end{pmatrix} \right) = \det \left( \begin{pmatrix} 2 & 1 & -1 \\ 0 & -\frac{3}{2} & \frac{5}{2} \\ 0 & \frac{3}{2} & \frac{3}{2} \end{pmatrix} \right) = \det \left( \begin{pmatrix} 2 & 1 & -1 \\ 0 & -\frac{3}{2} & \frac{5}{2} \\ 0 & 0 & 4 \end{pmatrix} \right) = 2 \cdot \left(-\frac{3}{2}\right) \cdot 4 = -12.$$

Die Regel von Sarrus ergibt, natürlich, das gleiche Ergebnis.  $\diamond$

Nun sei eine weitere Methode zur Berechnung der Determinante vorgestellt, der sogenannte Laplace'sche Entwicklungssatz. Dies hängt auch mit der Cramer'schen Regel zusammen.

**5.27 Definition** Sei  $A = (A_{n,m})_{n,m=1,\dots,N} \in \text{Mat}(N \times N, \mathbb{K})$  mit Spaltenvektoren  $a_n \in \mathbb{K}^N$ , d.h.  $A = (a_1, \dots, a_N)$ . Die Adjunkte von  $A$  ist die Matrix  $\hat{A} = (\hat{A}_{m,n})_{m,n=1,\dots,N} \in \text{Mat}(N \times N, \mathbb{K})$  definiert durch

$$\hat{A}_{m,n} = \det(a_1, \dots, a_{m-1}, e_n, a_{m+1}, \dots, a_N),$$

wobei auf der rechten Seite der  $m$ -te Spaltenvektor durch den  $n$ -ten Standardbasisvektor  $e_n \in \mathbb{K}^N$  ersetzt ist.

**5.28 Lemma** Sei  $\check{A}^{(n,m)} \in \text{Mat}((N-1) \times (N-1), \mathbb{K})$  die Matrix, die aus  $A$  durch Streichen der  $n$ -ten Zeile und der  $m$ -ten Spalte gewonnen wird. Dann gilt

$$\hat{A}_{m,n} = (-1)^{n+m} \det(\check{A}^{(n,m)}).$$

Die  $(N-1) \times (N-1)$  Matrix  $\check{A}^{(n,m)}$  heißt auch eine Untermatrix von  $A$ , und  $\det(\check{A}^{(n,m)})$  die zugehörige Unterdeterminante.

**Beweis.** Durch  $N - m$  benachbarte Transpositionen erhält man

$$\hat{A}_{m,n} = (-1)^{N-m} \det(a_1, \dots, a_{m-1}, a_{m+1}, \dots, a_N, e_n).$$

Nun werden  $N - n$  benachbarte Transpositionen der Zeilen durchgeführt (Erinnerung: die Determinante ändert sich nicht beim Transponieren der Matrix), so dass

$$\hat{A}_{m,n} = (-1)^{N-m} (-1)^{N-n} \det \left( \begin{pmatrix} \check{A}^{(n,m)} & 0 \\ b & 1 \end{pmatrix} \right),$$

wobei  $b$  ein Zeilenvektor der Größe  $N - 1$  ist, d.h.  $b^t \in \mathbb{K}^{N-1}$ , und die  $0 \in \mathbb{K}^{N-1}$ . Von dieser Blockform kann das Gewünschte direkt abgelesen werden, denn bei der Berechnung der Determinante tragen lediglich Permutationen  $\sigma \in S_N$  bei, für die  $\sigma(N) = N$  und die somit durch eine Permutation in  $S_{N-1}$  spezifiziert sind.  $\square$

**5.29 Satz** Sei  $A \in \text{Mat}(N \times N, \mathbb{K})$  mit Adjunker  $\widehat{A} \in \text{Mat}(N \times N, \mathbb{K})$ . Dann gilt

$$\widehat{A} A = \det(A) \mathbf{1}_N .$$

Falls  $\det(A) \neq 0$ , ist die Adjunkte im Wesentlichen die inverse Matrix:

$$A^{-1} = \frac{1}{\det(A)} \widehat{A} .$$

**Beweis.** Es werden die Einträge des Matrixproduktes berechnet:

$$\begin{aligned} (\widehat{A} A)_{m,n} &= \sum_{k=1}^N \widehat{A}_{m,k} A_{k,n} \\ &= \sum_{k=1}^N \det(a_1, \dots, a_{m-1}, e_k, a_{m+1}, \dots, a_N) A_{k,n} \\ &= \det(a_1, \dots, a_{m-1}, a_n, a_{m+1}, \dots, a_N) , \end{aligned} \tag{5.1}$$

wobei in der letzten Gleichung die Multilinearität von  $\det$  verwandt wurde. Nun ist die letzte Determinante nach Satz 5.19 immer gleich Null, es sei denn  $n = m$  und in diesem Fall ist sie  $\det(A)$ . Mit Hilfe des Kronecker Delta erhält man also

$$(\widehat{A} A)_{m,n} = \det(A) \delta_{n,m} ,$$

was genau die Behauptung ist. □

**5.30 Satz** (Laplace'scher Entwicklungssatz) Sei  $A \in \text{Mat}(N \times N, \mathbb{K})$  mit Unterdeterminanten  $\det(\check{A}^{(n,m)})$  definiert wie in Lemma 5.28. Dann gilt die Entwicklung der Determinante nach der  $m$ -ten Spalte

$$\det(A) = \sum_{k=1}^N A_{k,m} (-1)^{m+k} \det(\check{A}^{(k,m)}) ,$$

und der  $m$ -ten Zeile

$$\det(A) = \sum_{k=1}^N A_{m,k} (-1)^{m+k} \det(\check{A}^{(m,k)}) .$$

**Beweis.** Die Entwicklung nach der  $n$ -ten Spalte folgt aus der Gleichung (5.1) für  $n = m$ , wenn  $\widehat{A}_{m,k} = (-1)^{m+k} \det(\check{A}^{(k,m)})$  gemäß Lemma 5.28 eingesetzt wird. Mit  $\det(A) = \det(A^t)$  folgt die Entwicklung nach der  $m$ -ten Zeile aus der Entwicklung von  $\det(A^t)$  nach der  $m$ -ten Spalte. □

**5.31 Bemerkung** Der Laplace'sche Entwicklungssatz führt die Berechnung einer  $N \times N$  Determinante auf  $N$  Berechnungen von  $(N-1) \times (N-1)$  Determinanten zurück. Dies kann auch iterativ angewandt werden, so lange bis nur noch  $2 \times 2$  oder  $3 \times 3$  Determinanten zu berechnen sind. ◇

**5.32 Beispiel** Durch Entwicklung nach der dritten Spalte erhalten wir

$$\det \left( \begin{pmatrix} 2 & 1 & 0 & -1 \\ 3 & 0 & 0 & 1 \\ 5 & -1 & 2 & 7 \\ 1 & 2 & 0 & 1 \end{pmatrix} \right) = (-1)^{3+3} \cdot 2 \cdot \det \left( \begin{pmatrix} 2 & 1 & -1 \\ 3 & 0 & 1 \\ 1 & 2 & 1 \end{pmatrix} \right) .$$

Um die verbleibende  $3 \times 3$  Determinante zu berechnen, wird nach der zweiten Zeile entwickelt:

$$\det \left( \begin{pmatrix} 2 & 1 & -1 \\ 3 & 0 & 1 \\ 1 & 2 & 1 \end{pmatrix} \right) = -3 \cdot \det \left( \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} \right) - 1 \cdot \det \left( \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \right) = -3 \cdot 3 - 3 = -12.$$

Alternativ kann natürlich nach jeder Zeile oder Spalte entwickelt werden. Außerdem kann für die  $3 \times 3$  Matrix auch die Regel von Sarrus angewandt werden.  $\diamond$

Nun folgt noch eine Anwendung, die allerdings von wenig praktischer Bedeutung ist, denn das Lösen von linearen Gleichungssystemen mit Hilfe des Gauss-Algorithmus ist wesentlich effizienter.

**5.33 Satz** (Cramer'sche Regel) Sei  $A = (a_1, \dots, a_N) \in \text{Mat}(N \times N, \mathbb{K})$  mit Spaltenvektoren  $a_n \in \mathbb{K}^N$  invertierbar. Die Lösung  $x \in \mathbb{K}^N$  von  $Ax = b$  für gegebenes  $b \in \mathbb{K}^N$  ist gegeben durch

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}, \quad x_n = \frac{\det(a_1, \dots, a_{n-1}, b, a_{n+1}, \dots, a_N)}{\det(A)}.$$

**Beweis.** In der Tat, die  $n$ -te Komponente von  $x$  ist

$$x_n = (A^{-1}b)_n = \sum_{k=1}^N (A^{-1})_{n,k} b_k = \sum_{k=1}^N \frac{\det(a_1, \dots, a_{n-1}, e_k, a_{n+1}, \dots, a_N)}{\det(A)} b_k,$$

wobei im letzten Schritt die Formel für die Inverse aus Satz 5.29 verwandt wurde. Die Multilinearität von  $\det$  beendet den Beweis.  $\square$

**5.34 Beispiel** Seien

$$A = \begin{pmatrix} 2 & 1 & -1 \\ 3 & 0 & 1 \\ 1 & 2 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Schon oben wurde  $\det(A) = -12$  gezeigt. Nun kann die inverse Matrix durch Berechnung der Unterdeterminanten angegeben werden:

$$A^{-1} = \frac{-1}{12} \begin{pmatrix} -2 & -2 & 6 \\ -3 & 3 & -3 \\ 1 & -5 & -3 \end{pmatrix}^t = \frac{1}{12} \begin{pmatrix} 2 & 3 & -1 \\ 2 & -3 & 5 \\ -6 & 3 & 3 \end{pmatrix}.$$

Die erste Komponente des Lösungsvektors  $x$  von  $Ax = b$  ist

$$x_1 = \frac{-1}{12} \det \left( \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & 1 \\ 3 & 2 & 1 \end{pmatrix} \right) = \frac{-1}{12} (3 - 4 - 2 - 2) = \frac{5}{12}.$$

Dies kann auch durch Matrix-Vektor-Multiplikation aus  $A^{-1}b$  berechnet werden.  $\diamond$



## 6 Eigenwerte und Eigenvektoren

**6.1 Definition** Sei  $T \in \mathcal{L}(V)$  eine lineare Abbildung auf einem Vektorraum  $V$  über  $\mathbb{K}$ . Ein Skalar  $z \in \mathbb{K}$  heißt Eigenwert von  $T$  genau dann, wenn es einen sogenannten Eigenvektor  $v \in V, v \neq 0$ , gibt mit

$$T v = z v .$$

Diese Gleichung heißt eine Eigenwertgleichung.

**6.2 Satz** Sei  $S \in \mathcal{L}(V, W)$  eine invertierbare lineare Abbildung zwischen Vektorräumen  $V$  und  $W$  über  $\mathbb{K}$ . Dann gilt für  $T \in \mathcal{L}(V)$ :

$$z \in \mathbb{K} \text{ Eigenwert von } T \iff z \in \mathbb{K} \text{ Eigenwert von } STS^{-1} .$$

Insbesondere, wenn  $V$  endlich dimensional ist mit geordneter Basis  $\mathcal{B}$ , dann gilt für die darstellende Matrix  $T_{\mathcal{B}}^{\mathcal{B}}$ :

$$z \in \mathbb{K} \text{ Eigenwert von } T \iff z \in \mathbb{K} \text{ Eigenwert von } T_{\mathcal{B}}^{\mathcal{B}} .$$

**Beweis.** "  $\implies$  " Sei  $T v = z v$ . Dann gilt nach Einfügen von  $\mathbf{1}_V = S^{-1} S$  auch  $TS^{-1} S v = z v$  und somit auch

$$STS^{-1}(Sv) = z Sv .$$

Also ist  $Sv \in W$  Eigenvektor von  $STS^{-1} \in \mathcal{L}(W)$  zum Eigenwert  $z$ . Diese Argumentation kann umgekehrt werden, so dass auch "  $\impliedby$  " folgt. Der Zusatz ist der Spezialfall, in dem  $S = I_{\mathcal{B}} : V \rightarrow \mathbb{K}^N$  mit  $N = \dim(V)$  ist.  $\square$

Somit ist es ausreichend, Techniken zur Berechnung von Eigenwerten von Matrizen zu entwickeln. Der wichtigste Sachverhalt in dieser Hinsicht ist der folgende.

**6.3 Satz** Sei  $A \in \text{Mat}(N \times N, \mathbb{K})$ . Dann gilt

$$z \in \mathbb{K} \text{ Eigenwert von } A \iff \det(A - z \mathbf{1}_N) = 0 .$$

Die Eigenwerte von  $A$  sind also genau die Menge der Nullstellen in  $\mathbb{K}$  des Polynoms

$$z \in \mathbb{K} \mapsto p_A(z) = \det(A - z \mathbf{1}_N) .$$

Es heißt charakteristisches Polynom von  $A$  und ist  $N$ -ten Grades. Es erfüllt

$$\det(A - z \mathbf{1}_N) = (-z)^N + \text{Tr}(A) (-z)^{N-1} + q(z) + \det(A) ,$$

wobei  $q = \sum_{n=1}^{N-2} q_n z^n$  ein Polynom vom Grade höchstens  $N - 2$  ohne konstanten Term ist, und  $\text{Tr}(A)$  die Spur von  $A$  ist, welche definiert ist als die Summe der Diagonaleinträge von  $A$ :

$$\text{Tr}(A) = \sum_{n=1}^N A_{n,n} .$$

**Beweis.** Die Eigenwertgleichung  $Av = zv$  für  $v \in \mathbb{K}^N$  kann umgeschrieben werden als  $(A - z \mathbf{1}_N)v = 0$ . Nun ist die Existenz eines Vektors  $v \in \mathbb{K}^N, v \neq 0$ , äquivalent dazu, dass die Matrix  $A - z \mathbf{1}_N$  nicht invertierbar ist. Letzteres ist nach Satz 5.21 äquivalent dazu, dass die Determinante der Matrix  $A - z \mathbf{1}_N$

verschwindet. Insbesondere sei beachtet, dass  $\det(A - z \mathbf{1}_N) = 0$  auch impliziert, dass ein Eigenvektor  $v$  zum Eigenwert  $z$  existiert.

Nun soll noch die Determinante in  $z$  entwickelt werden. Hierzu schreiben wir die Definition aus:

$$\det(A - z \mathbf{1}_N) = \sum_{\sigma \in S_N} \operatorname{sgn}(\sigma) \prod_{n=1}^N (A_{n,\sigma(n)} - z \delta_{n,\sigma(n)}),$$

wobei  $\delta_{n,m}$  das Kronecker Delta ist. Hieraus ist sofort ersichtlich, dass der Grad höchstens  $N$  ist. Außerdem sind die Potenzen  $z^N$  und  $z^{N-1}$  nur in dem Summanden  $\sigma = \operatorname{id}$  enthalten, denn für  $\sigma \neq \operatorname{id}$  ist der Faktor ein Polynom in  $z$  vom Grad höchstens  $N - 2$ . Also

$$\det(A - z \mathbf{1}_N) = \left( \prod_{n=1}^N (A_{n,n} - z) \right) + q'(z) = (-z)^N + \operatorname{Tr}(A)(-z)^{N-1} + q'(z),$$

wobei  $q'$  und  $q''$  Polynome vom Grad höchstens  $N - 2$  sind. Der konstante Term kann berechnet werden, indem  $z = 0$  gesetzt wird. Dies zeigt, dass er  $\det(A)$  ist.  $\square$

**6.4 Bemerkung** Wenn  $T \in \mathcal{L}(V)$  eine lineare Abbildung auf einem  $N$ -dimensionalen Vektorraum  $V$  ist, so ist sein charakteristisches Polynom

$$p_T(z) = \det(T_{\mathcal{B}}^{\mathcal{B}} - z \mathbf{1}_N)$$

unabhängig von der Wahl der Basis  $\mathcal{B}$  von  $V$ . In der Tat, wenn  $\mathcal{C}$  eine andere Basis ist, so gilt

$$\begin{aligned} \det(T_{\mathcal{C}}^{\mathcal{C}} - z \mathbf{1}_N) &= \det((\mathbf{1}_V)_{\mathcal{C}}^{\mathcal{B}} T_{\mathcal{B}}^{\mathcal{B}} ((\mathbf{1}_V)_{\mathcal{C}}^{\mathcal{B}})^{-1} - z (\mathbf{1}_V)_{\mathcal{C}}^{\mathcal{C}} ((\mathbf{1}_V)_{\mathcal{C}}^{\mathcal{C}})^{-1}) \\ &= \det((\mathbf{1}_V)_{\mathcal{C}}^{\mathcal{B}} (T_{\mathcal{B}}^{\mathcal{B}} - z \mathbf{1}_V) ((\mathbf{1}_V)_{\mathcal{C}}^{\mathcal{B}})^{-1}) = \det(T_{\mathcal{B}}^{\mathcal{B}} - z \mathbf{1}_N), \end{aligned}$$

wobei im letzten Schritt die Rechenregeln der Determinante verwandt wurden. Die gleiche Rechnung zeigt auch, dass wenn  $A$  und  $B$  zwei ähnliche  $N \times N$  Matrizen sind (d.h. es gibt ein invertierbares  $M$  mit  $A = MBM^{-1}$ ), dann gilt  $p_A = p_B$ . Die Umkehrung ist allerdings falsch, denn es gibt Beispiele von Matrizen mit gleichem charakteristischem Polynom, die nicht ähnlich sind, z.B.

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

In der Tat, die Einheitsmatrix  $A$  ist nur ähnlich zu sich selbst, aber dennoch stimmen die beiden charakteristischen Polynome überein.  $\diamond$

**6.5 Beispiel** Sei

$$A = \begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix}.$$

Also

$$p_A(z) = (5 - z)^2 - 1 = z^2 - 10z + 24 = (z - 4)(z - 6).$$

Also sind die Eigenwerte  $z_1 = 4$  und  $z_2 = 6$ .  $\diamond$

**6.6 Beispiel** Für  $2 \times 2$  Matrizen ist es einfach, eine allgemeine Formel für die Eigenwerte anzugeben. Sei

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Dann ist das charakteristische Polynom

$$p_A(z) = \det \left( \begin{pmatrix} a-z & b \\ c & d-z \end{pmatrix} \right) = (a-z)(d-z) - bc = z^2 - (a+d)z + (ad-bc).$$

Dies kann auch geschrieben werden als

$$p_A(z) = z^2 - \operatorname{Tr}(A)z + \det(A),$$

was auch aus Satz 6.3 hervorgeht, weil  $q = 0$  für  $N = 2$ . Die Nullstellen sind

$$z_{\pm} = \frac{\operatorname{Tr}(A)}{2} \pm \sqrt{\frac{\operatorname{Tr}(A)^2}{4} - \det(A)}.$$

Nun ist es möglich, dass  $A$  reell ist und dann als Abbildung auf dem reellen Vektorraum  $\mathbb{R}^2$  aufgefasst wird, das charakteristische Polynom aber nur komplexe Nullstellen hat. Z.B.

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \implies z_{\pm} = \pm i.$$

Dann hat  $A$  keinen Eigenwert (weil die Wurzeln nicht reell sind). Das gleich gilt, wenn  $A$  aufgefasst wird als lineare Abbildung auf dem reellen Vektorraum  $\mathbb{C}^2$ . Wenn  $A$  hingegen als lineare Abbildung auf dem komplexen Vektorraum  $\mathbb{C}^2$  aufgefasst wird, so hat  $A$  zwei Eigenwerte  $z_{\pm}$ . Diese Eigenwerte sind komplex konjugiert zueinander.  $\diamond$

**6.7 Bemerkung (Wichtig!)** Es gibt lineare Abbildungen auf reellen Vektorräumen ohne Eigenwerte. Auf komplexen Vektorräumen hingegen ist die Situation ganz anders, denn der folgende Satz zeigt, dass die Anzahl der Eigenwerte gleich der Dimension des Vektorraumes ist, vorausgesetzt die Eigenwerte werden mit ihrer algebraischen Multiplizität gezählt.  $\diamond$

**6.8 Satz** Sei  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ , und  $V$  ein  $N$ -dimensionaler Vektorraum über  $\mathbb{K}$  und  $T \in \mathcal{L}(V)$ . Dann besitzt das charakteristische Polynom eine Faktorisierung

$$p_T(z) = (-1)^N \prod_{k=1}^K (z - z_k)^{\alpha(z_k)}, \quad (6.1)$$

mit paarweise verschiedenen  $z_k \in \mathbb{C}$  und mit  $\alpha(z_k) \in \mathbb{N}$ , die folgende Gleichung erfüllen:

$$\sum_{k=1}^K \alpha(z_k) = N. \quad (6.2)$$

Falls  $z_k \in \mathbb{K}$ , so ist  $z_k$  Eigenwert von  $T$  und dann heißt die Zahl  $\alpha(z_k)$  auch die algebraische Vielfachheit oder algebraische Multiplizität des Eigenwertes  $z_k$ . Insbesondere hat  $T$  höchstens  $N$  verschiedene Eigenwerte.

**Beweis.** Nach dem Fundamentalsatz der Algebra (Satz 1.30) existieren  $N$  komplexe Nullstellen  $z'_1, \dots, z'_N \in \mathbb{C}$  gibt, so dass

$$p_A(z) = (-1)^N (z - z'_1) \cdots (z - z'_N).$$

Zusammenfassen gleicher  $z'_n$ 's führt schon auf die Darstellung im Satz. Kombiniert mit Satz 6.3 impliziert dies die Aussage über die Eigenwerte.  $\square$

### 6.9 Beispiel Sei

$$A = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Dann ist das charakteristische Polynom berechnet mit Hilfe der Regel von Sarrus

$$p_A(z) = \det \left( \begin{pmatrix} 1-z & 2 & 0 \\ -1 & -z & 1 \\ 1 & 0 & -z \end{pmatrix} \right) = (1-z)z^2 + 2 - (-z) \cdot 2 \cdot (-1) = (1-z)(z^2 + 2).$$

Also sind die Nullstellen von  $p_A$  gegeben durch

$$z_1 = 1, \quad z_2 = i\sqrt{2}, \quad z_3 = -i\sqrt{2}.$$

Wenn nun  $A$  als lineare Abbildung auf  $\mathbb{R}^3$  aufgefasst wird, so ist lediglich  $z_1 = 1$  ein Eigenwert, weil nämlich  $z_2$  und  $z_3$  nicht in  $\mathbb{R}$  liegen. Wird hingegen  $A$  als lineare Abbildung auf  $\mathbb{C}^3$  aufgefasst (welche als Zusatzeigenschaft nur reelle Einträge hat), so sind alle drei Nullstellen auch Eigenwerte. Wiederum sind die komplexen Eigenwerte zueinander konjugiert.  $\diamond$

**6.10 Bemerkung** Das charakteristische Polynom in Beispiel 6.9 ist  $p_A(z) = (1-z)(z^2 + 2)$ . Es kann über  $\mathbb{R}$ , d.h. als Polynom in  $\mathbb{R}[z]$ , nicht in Linearfaktoren faktorisiert werden. Nach dem Fundamentalsatz der Algebra kann aber jedes Polynom in  $\mathbb{C}[z]$  in Linearfaktoren faktorisiert werden.  $\diamond$

Wenn ein Eigenwert  $z$  von  $T$  bestimmt ist (z.B. durch Bestimmen einer Nullstelle des charakteristischen Polynomes), so können Eigenvektoren zu  $z$  als Lösung des homogenen linearen Gleichungssystems  $(T - z \mathbf{1})v = 0$  berechnet werden (es existieren auf jeden Fall solche Lösungen, da ja  $T - z \mathbf{1}$  nicht invertierbar ist). Dies kann mit Hilfe des Gauss Algorithmus durchgeführt werden.

### 6.11 Beispiel Sei

$$A = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},$$

wie in Beispiel 6.9. Nach obiger Rechnung sind die Nullstellen des charakteristischen Polynomes  $z_1 = 1$ ,  $z_2 = i\sqrt{2}$  und  $z_3 = -i\sqrt{2}$ . Um Eigenvektoren  $v_1$  zu  $z_1$  zu berechnen, benötigen wir

$$A - z_1 \mathbf{1} = \begin{pmatrix} 0 & 2 & 0 \\ -1 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix}.$$

Der Kern wird mit dem Gauss-Algorithmus berechnet (erst Vertauschen erster und zweiter Zeile, dann algorithmisch):

$$\begin{array}{ccc|c} -1 & -1 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array} \longrightarrow \begin{array}{ccc|c} -1 & -1 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}.$$

Also ist die Menge aller Eigenvektoren zum Eigenwert  $z_1$ :

$$\text{Ker}(A - z_1 \mathbf{1}) = \text{span}(v_1), \quad v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Der Span kann entweder reell oder komplex gewählt werden, je nachdem, ob  $A$  als Abbildung auf dem  $\mathbb{R}^3$  oder  $\mathbb{C}^3$  aufgefasst wird. Analog findet man  $\text{Ker}(A - z_2 \mathbf{1}) = \text{span}(v_2)$  und  $\text{Ker}(A - z_3 \mathbf{1}) = \text{span}(v_3)$

$$v_2 = \begin{pmatrix} -2 \\ 1 - i\sqrt{2} \\ i\sqrt{2} \end{pmatrix}, \quad v_3 = \begin{pmatrix} -2 \\ 1 + i\sqrt{2} \\ -i\sqrt{2} \end{pmatrix}.$$

Dies sind jedoch offensichtlich nur im komplexen Fall Eigenvektoren. ◇

Dieses Beispiel illustriert auch folgenden Sachverhalt.

**6.12 Satz** Sei  $A \in \text{Mat}(N \times N, \mathbb{R})$  eine reelle Matrix, die aufgefasst wird als lineare Abbildung auf dem komplexen Vektorraum  $\mathbb{C}^N$ . Wenn  $z \in \mathbb{C}$  Eigenwert ist, dann ist auch sein komplex Konjugiertes  $\bar{z} \in \mathbb{C}$  Eigenwert von  $A$  und die Eigenvektoren sind auch komplex konjugiert zueinander.

**Beweis.** Sei  $Av = zv$ . Dann ergibt komplexes Konjugieren aller Matrix- und Vektoreinträge sowie von  $z$ , dass  $A\bar{z} = \bar{z}\bar{v}$ , weil ja  $A$  reell ist. Hieraus folgen alle Aussagen. □

Nun folgen wichtige strukturelle Aussagen über Eigenvektoren.

**6.13 Satz** Für jede lineare Abbildung  $T \in \mathcal{L}(V)$  sind Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig, d.h. wenn  $v_1, \dots, v_K$  Eigenvektoren zu paarweise verschiedenen Eigenwerten  $z_1, \dots, z_K$  von  $T$  sind, so sind  $v_1, \dots, v_K$  linear unabhängig.

**Beweis.** Der Beweis wird durch Induktion über  $K$  geführt. Für  $K = 1$  ist die Aussage richtig, da  $v_1 \neq 0$ . Nun zum Induktionsschritt von  $K - 1$  zu  $K$ . Seien  $\lambda_1, \dots, \lambda_K \in \mathbb{K}$ , so dass

$$\sum_{k=1}^K \lambda_k v_k = 0.$$

Multiplikation dieser Gleichung mit  $z_K$  und Anwenden von  $T$  liefert:

$$\sum_{k=1}^K \lambda_k z_K v_k = 0, \quad \sum_{k=1}^K \lambda_k z_k v_k = 0.$$

Somit ergibt Subtraktion der beiden Gleichungen

$$\sum_{k=1}^{K-1} \lambda_k (z_K - z_k) v_k = 0.$$

Da  $z_K - z_k \neq 0$ , ergibt die Induktionshypothese, dass  $\lambda_k = 0$  für  $k = 1, \dots, K - 1$ . Wird dies in  $\sum_{k=1}^K \lambda_k v_k = 0$  eingesetzt, so folgt auch  $\lambda_K = 0$ . □

**6.14 Definition** Sei  $T \in \mathcal{L}(V)$  eine lineare Abbildung auf einem Vektorraum  $V$  über  $\mathbb{K}$  und sei  $z \in \mathbb{K}$  ein Eigenwert von  $T$ . Der Eigenraum  $E_T(z)$  ist der Kern von  $T - z \mathbf{1}$ . Seine Dimension  $\beta(z)$  heißt die geometrische Vielfachheit oder geometrische Multiplizität von  $z$  als Eigenwert von  $T$ .

**6.15 Beispiel** Für  $A$  wie in Beispiel 6.11 wirkend auf  $\mathbb{C}^3$ , ist  $E_A(z_j) = \text{span}(v_j)$  mit den dort angegebenen Vektoren  $v_j$ ,  $j = 1, 2, 3$ . Außerdem  $\alpha(z_j) = \beta(z_j) = 1$ . ◇

**6.16 Satz** Sei  $T \in \mathcal{L}(V)$  auf einem Vektorraum  $V$  über  $\mathbb{K}$ . Dann gilt  $1 \leq \beta(z_k) \leq \alpha(z_k)$  für alle Eigenwerte  $z_k \in \mathbb{K}$ .

**Beweis.** Sei  $z_k$  Eigenwert von  $T$  mit geometrischer Multiplizität  $\beta(z_k)$ . Dann sei  $v_1, \dots, v_{\beta(z_k)} \in V$  eine Basis des Eigenraumes  $E_T(z_k)$ . Diese Basis wird vervollständigt zu einer geordneten Basis  $\mathcal{B} = (v_1, \dots, v_N)$  von  $V$ . Da das Bild von  $E_T(z_k)$  unter  $T$  wieder in  $E_T(z_k)$  ist und hierauf die Identität ist, ist die darstellende Matrix von der Form

$$T_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} z_k \mathbf{1}_{\beta(z_k)} & B \\ 0 & C \end{pmatrix},$$

wobei  $B$  und  $C$  Matrizen geeigneter Größe sind (genau ist  $C$  eine  $(N - \beta(z_k)) \times (N - \beta(z_k))$  Matrix), und  $B$  eine  $\beta(z_k) \times (N - \beta(z_k))$  Matrix). Somit ist auch  $T_{\mathcal{B}}^{\mathcal{B}} - z \mathbf{1}_N$  von dieser Blockstruktur und somit folgt, weil die Determinante einer  $2 \times 2$  Blockmatrix mit einer verschwindenden Nebendiagonalen das Produkt der Determinanten der Diagonalblöcke ist,

$$p_T(z) = \det(T_{\mathcal{B}}^{\mathcal{B}} - z \mathbf{1}_N) = \det \left( \begin{pmatrix} (z_k - z) \mathbf{1}_{\beta(z_k)} & B \\ 0 & C - z \mathbf{1} \end{pmatrix} \right) = (z_k - z)^{\beta(z_k)} \det(C - z \mathbf{1}).$$

Also folgt in der Tat  $\alpha(z_k) \geq \beta(z_k)$ . □

**6.17 Definition**  $T \in \mathcal{L}(V)$  heißt diagonalisierbar, genau dann wenn es eine Basis  $\mathcal{B}$  von  $V$  bestehend aus Eigenvektoren von  $T$  gibt.

**6.18 Satz** Sei  $T \in \mathcal{L}(V)$  auf einem Vektorraum  $V$  über  $\mathbb{K}$  der Dimension  $N < \infty$ . Dann sind äquivalent:

- (i)  $T$  diagonalisierbar.
- (ii)  $\alpha(z_k) = \beta(z_k)$  für alle Eigenwerte  $z_k \in \mathbb{K}$  und  $\sum_k \alpha(z_k) = N$ , wobei die Summe über alle Eigenwerte  $z_k \in \mathbb{K}$  läuft, anders als in (6.2), wo sie über alle Nullstellen in  $\mathbb{C}$  des charakteristischen Polynomes läuft.
- (iii)  $\sum_k \beta(z_k) = N$ , wobei die Summe über alle Eigenwerte  $z_k \in \mathbb{K}$  läuft.

Falls  $T$  diagonalisierbar ist, so ist die darstellende Matrix in der geordneten Basis  $\mathcal{B} = (v_1, \dots, v_N)$  von Eigenvektoren zu den Eigenwerten  $z'_1, \dots, z'_N$ , aufgelistet mit ihrer Multiplizität, gegeben durch eine Diagonalmatrix:

$$T_{\mathcal{B}}^{\mathcal{B}} = (z'_m \delta_{n,m})_{n,m=1,\dots,N} = \begin{pmatrix} z'_1 & & & \\ & z'_2 & & \\ & & \ddots & \\ & & & z'_N \end{pmatrix}. \quad (6.3)$$

**Beweis.** (i)  $\implies$  (ii) Wenn es eine Basis von Eigenvektoren gibt, so gilt  $\sum_k \beta(z_k) = N$ , wobei die Summe über alle Eigenwerte  $z_k \in \mathbb{K}$  läuft. Da  $\sum_k \alpha(z_k) \leq N$  immer gilt (mit der Summe über die Eigenwerte) und  $\beta(z_k) \leq \alpha(z_k)$  nach Satz 6.16, folgen alle Behauptungen. (ii)  $\implies$  (iii) ist offensichtlich. (iii)  $\implies$  (i) folgt aus Satz 6.13.

Nun zum letzten Punkt. Es sei daran erinnert, dass  $T_{\mathcal{B}}^{\mathcal{B}} = I_{\mathcal{B}} T (I_{\mathcal{B}})^{-1}$  mit  $(I_{\mathcal{B}})^{-1} = (v_1, \dots, v_N)$ . Also

$$(T_{\mathcal{B}}^{\mathcal{B}})_{n,m} = (I_{\mathcal{B}} T (I_{\mathcal{B}})^{-1} e_m)_n = (I_{\mathcal{B}} T v_m)_n = (I_{\mathcal{B}} z'_m v_m)_n = z'_m (I_{\mathcal{B}} (I_{\mathcal{B}})^{-1} e_m)_n = z'_m \delta_{n,m},$$

wobei  $e_m$  den  $m$ -ten Standardbasisvektor bezeichnet. □

**6.19 Bemerkung** Für eine Matrix  $A$  bedeutet die Diagonalisierbarkeit, dass wenn  $M^{-1} = (v_1, \dots, v_N)$  gebildet wird aus einer Basis von Eigenvektoren, dann ist  $D = MAM^{-1}$  diagonal, mit den Eigenwerten aus Diagonaleinträgen. In der Tat,  $A$  ist eine lineare Abbildung auf  $\mathbb{K}^N$ . Wenn nun  $\mathcal{B} = (v_1, \dots, v_N)$  die geordnete Basis von Eigenvektoren ist, dann ist ja gerade  $A_{\mathcal{B}}^{\mathcal{B}} = MAM^{-1}$ .  $\diamond$

**6.20 Beispiel** Sei  $A$  die  $3 \times 3$  Matrix aus Beispiel 6.11, wo auch deren Eigenwerte und Eigenvektoren berechnet wurden. Dann setze

$$M^{-1} = \begin{pmatrix} 1 & -2 & -2 \\ 0 & 1 - i\sqrt{2} & 1 + i\sqrt{2} \\ 1 & i\sqrt{2} & -i\sqrt{2} \end{pmatrix}.$$

Es gilt dann nach Satz 6.18

$$MAM^{-1} = \begin{pmatrix} 1 & & 0 \\ 0 & i\sqrt{2} & 0 \\ 0 & 0 & -i\sqrt{2} \end{pmatrix}.$$

Dies kann auch explizit verifiziert werden.  $\diamond$

Als eine Anwendung der Diagonalisierung einer diagonalisierbaren linearen Abbildung sollen nun Funktionen  $f$  von linearen Abbildungen betrachtet und berechnet werden, was z.B. zum Lösen von linearen Differentialgleichungssystemen sehr hilfreich ist. Wir schränken uns hierbei auf sogenannte ganze Potenzreihen ein (siehe Analysis), d.h.  $f$  ist von der Gestalt

$$f(z) = \sum_{k=0}^{\infty} f_k z^k, \tag{6.4}$$

wobei  $f_n \in \mathbb{C}$  die Koeffizienten sind und die Reihe für alle  $z \in \mathbb{C}$  konvergiert (Konvergenzradius unendlich). Ein Spezialfall sind Polynome und selbst für solche sind folgende Definition und folgender Satz von Interesse.

**6.21 Definition** Sei  $T \in \mathcal{L}(V)$  auf einem endlich dimensionalen Vektorraum  $V$ . Für eine ganze Potenzreihe  $f$  der Form (6.4), wird eine neue lineare Abbildung  $f(T) \in \mathcal{L}(V)$  definiert durch die konvergente Reihe

$$f(T) = \sum_{k=0}^{\infty} f_k T^k.$$

**6.22 Satz** Sei  $T$  diagonalisierbar und  $\mathcal{B}$  eine geordnete Basis von Eigenvektoren, so dass (6.3) gilt. Dann

$$f(T)_{\mathcal{B}}^{\mathcal{B}} = (f(z'_m) \delta_{n,m})_{n,m=1,\dots,N} = \begin{pmatrix} f(z'_1) & & & \\ & f(z'_2) & & \\ & & \ddots & \\ & & & f(z'_N) \end{pmatrix}.$$

**Beweis.** Sei  $D = f(T)_{\mathcal{B}}^{\mathcal{B}}$  die in (6.3) gegebene Diagonalmatrix. Wir beginnen mit dem Monom  $f(z) = z^k$ . Dann ist  $f(T) = T^k$ . Somit folgt nach mehrfachen Einfügen von  $(I_{\mathcal{B}})^{-1} I_{\mathcal{B}} = \mathbf{1}_V$

$$f(T)_{\mathcal{B}}^{\mathcal{B}} = I_{\mathcal{B}} T^k (I_{\mathcal{B}})^{-1} = I_{\mathcal{B}} T (I_{\mathcal{B}})^{-1} I_{\mathcal{B}} T (I_{\mathcal{B}})^{-1} \dots I_{\mathcal{B}} T (I_{\mathcal{B}})^{-1} = D^k.$$

Für eine beliebige ganze Potenzreihe gilt also (da  $I_B$  in die Summe hereingezogen werden darf, wie durch Restabschätzung verifiziert werden kann)

$$f(T)_B^B = I_B \sum_{k=0}^{\infty} f_k T^k (I_B)^{-1} = \sum_{k=0}^{\infty} f_k I_B T^k (I_B)^{-1} = \sum_{k=0}^{\infty} f_k D^k = f(D).$$

Dies ist genau die Aussage, denn  $f(D)$  ist die angegebene Matrix. □

**6.23 Bemerkung** Für eine diagonalisierbare Matrix  $A$  mit Diagonalmatrix  $D = MAM^{-1}$ , bedeutet dies  $Mf(A)M^{-1} = f(D)$ . Hierbei ist  $f(D)$  eine Diagonalmatrix mit Diagonaleinträgen  $f(z'_n)$ . Umschreiben zeigt also  $f(A) = M^{-1}f(D)M$ . Natürlich könnte, z.B. für ein Polynom  $f(A)$  auch durch Matrixmultiplikationen explizit berechnet werden, aber dies ist wesentlich mühsamer als die Verwendung der Diagonalisierung. Diese Methode des Vorgehens wird auch mit Spektralkalkül bezeichnet, weil lediglich die Diagonalmatrix verändert wird, die ja genau die Eigenwerte auf der Diagonale enthält, und diese Eigenwerte machen das Spektrum aus (siehe Definition 6.32 weiter unten). ◇

**6.24 Beispiel** Sei wieder  $A$  die  $3 \times 3$  Matrix aus Beispiel 6.11, die in Beispiel 6.20 diagonalisiert wurde mit Hilfe der Matrix  $M$  diagonalisiert wurde zu  $D = MAM^{-1}$ . Es soll nun  $e^A$  berechnet werden, d.h.  $f(z) = e^z$  ist die Exponentialfunktion. Es gilt

$$f(A) = M^{-1}f(D)M = M^{-1} \begin{pmatrix} e & 0 & 0 \\ 0 & e^{i\sqrt{2}} & 0 \\ 0 & 0 & e^{-i\sqrt{2}} \end{pmatrix} M.$$

Nun wurde  $M^{-1}$  schon oben berechnet. Hieraus kann  $M$  durch Invertieren berechnet werden (mit dem erweiterten Gauss Algorithmus, oder über die Berechnung der Adjunkten). Danach muss lediglich das Produkt von 3 Matrizen berechnet werden. ◇

**6.25 Bemerkung** Nach Satz 6.8 ist jedes charakteristische Polynom über den komplexen Zahlen in Linearfaktoren faktorisiert (dies folgte direkt aus dem Fundamentalsatz der Algebra). Somit gilt für die algebraischen Multiplizitäten einer linearen Abbildung auf einem komplexen Vektorraum  $\sum_k \alpha(z_k) = N$ , wobei die Summe über alle Eigenwerte  $z_k \in \mathbb{C}$  läuft. Dies impliziert allerdings nicht, dass diese Abbildung diagonalisierbar ist. Hier ist ein explizites Beispiel einer Matrix  $A \in \text{Mat}(N \times N, \mathbb{C})$ , die nicht diagonalisierbar ist:

$$A = (z_1 \delta_{n,m} + \delta_{n+1,m})_{n,m=1,\dots,N} = \begin{pmatrix} z_1 & 1 & 0 & 0 \\ 0 & z_1 & 1 & 0 \\ & \ddots & \ddots & \ddots \\ & & 0 & z_1 & 1 \\ 0 & & & 0 & z_1 \end{pmatrix}.$$

Diese Matrix ist eine sogenannte obere Dreiecksmatrix, denn unter der Diagonalen steht nur der Eintrag Null. Somit ist ihr charakteristisches Polynom sehr einfach zu berechnen als  $p_A(z) = (z_1 - z)^N$ . Insbesondere faktorisiert dieses Polynom und die algebraische Multiplizität des einzigen Eigenwertes  $z_1$  ist  $\alpha(z_1) = N$ . Dennoch gibt es nur einen einzigen linear unabhängigen Eigenvektor, da nämlich

$$\text{Ker}(A - z_0 \mathbf{1}) = \text{Ker} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ & \ddots & \ddots & \ddots \\ & & 0 & 1 \\ 0 & & & 0 \end{pmatrix} = \text{span}(e_1).$$



Also ist die geometrische Multiplizität  $\beta(z_1) = 1$ . Eine Matrix von der Gestalt von  $A$  (konstante Diagonale und obere Nebendiagonale mit Einträgen 1, sonst 0) heißt ein Jordan Block. Der Satz von der Jordan'schen Normalform besagt, dass jede lineare Abbildung in einer geeigneten Basis eine blockdiagonale Matrix mit Jordan Blöcken auf der Diagonale ist. Dieser Satz wird im kommenden Semester bewiesen. Das folgende Ergebnis ist eine Vorstufe dazu.  $\diamond$

**6.26 Definition** Eine lineare Abbildung  $T \in \mathcal{L}(V)$  heißt trigonalisierbar genau dann, wenn es eine geordnete Basis  $\mathcal{B}$  von  $V$  gibt, so dass  $T_{\mathcal{B}}^{\mathcal{B}}$  eine obere Dreiecksmatrix ist. Insbesondere heißt eine Matrix  $B \in \text{Mat}(N \times N, \mathbb{K})$  trigonalisierbar genau dann, wenn es ein invertierbares  $M \in \text{Mat}(N \times N, \mathbb{K})$  gibt, so dass  $MBM^{-1}$  eine obere Dreiecksmatrix ist.

**6.27 Bemerkung** Die Triangulierung von  $A$  ist nicht die obere Dreiecksmatrix, die durch den Gauss-Algorithmus berechnet wird. In der Tat, dort wird ein  $M$  dergestalt bestimmt, dass  $MA$  eine obere Dreiecksmatrix ist, und nicht  $MAM^{-1}$ .  $\diamond$

**6.28 Satz** Sei  $T \in \mathcal{L}(V)$  auf einem Vektorraum  $V$  über  $\mathbb{K}$  mit endlicher Dimension  $N = \dim(V) < \infty$ . Dann sind äquivalent:

- (i)  $T$  ist trigonalisierbar.
- (ii) Das charakteristische Polynom faktorisiert in  $\mathbb{K}$  (d.h.  $\sum_k \alpha(z_k) = N$ , wobei die Summe über alle Eigenwerte  $z_k \in \mathbb{K}$  läuft).
- (iii) Es gibt eine Fahne für  $T$ , d.h. eine Folge von Unterräumen

$$\{0\} = V_0 \subset V_1 \subset V_2 \subset \dots \subset V_{N-1} \subset V_N = V$$

mit  $\dim(V_n) = n$  und  $TV_n \subset V_n$  (d.h.  $V_n$  ist  $T$ -invariant).

**Beweis.** (i) $\implies$ (ii) Das charakteristische Polynom  $p_T(z) = \det(T_{\mathcal{B}}^{\mathcal{B}} - z\mathbf{1}_N)$  ist die Determinante einer oberen Dreiecksmatrix und somit gegeben durch das Produkt der Diagonaleinträge von  $T_{\mathcal{B}}^{\mathcal{B}} - z\mathbf{1}_N$ . Also faktorisiert es.

(ii) $\implies$ (iii) Es reicht, die Existenz der Fahne für die darstellende Matrix  $A = T_{\mathcal{B}}^{\mathcal{B}}$  nachzuweisen, denn die Fahne von  $T$  erhält man dann durch Abbilden mit der Koordinatenabbildung  $(I_{\mathcal{B}})^{-1}$ . Für  $A$  zeigen wir die Existenz der Fahne durch Induktion über  $N$ . Der Induktionsanfang  $N = 1$  ist offensichtlich. Nun zum Induktionsschritt  $N - 1 \rightarrow N$ . Sei  $z_1 \in \mathbb{K}$  ein Eigenwert von  $A$  mit Eigenvektor  $v_1 \in \mathbb{K}^N$  (dieser Eigenwert existiert nach Voraussetzung). Dann wird  $v_1$  zu einer Basis von  $\mathcal{B} = (v_1, \dots, v_N)$  von  $\mathbb{K}^N$  vervollständigt. Es gilt dann mit  $M^{-1} = (v_1, \dots, v_N)$  (ähnlich wie im Beweis von Satz 6.16)

$$MAM^{-1} = \begin{pmatrix} z_1 & B \\ 0 & C \end{pmatrix},$$

wobei  $C \in \text{Mat}((N - 1) \times (N - 1), \mathbb{K})$ . Für diese Abbildung auf dem  $\mathbb{K}^{N-1}$  gibt es nach Induktionsvoraussetzung eine invertierbares  $K$ , so dass  $KCK^{-1}$  eine obere Dreiecksmatrix ist. Also ist

$$\begin{pmatrix} 1 & 0 \\ 0 & K \end{pmatrix} MAM^{-1} \begin{pmatrix} 1 & 0 \\ 0 & K \end{pmatrix}^{-1} = \begin{pmatrix} z_1 & BK^{-1} \\ 0 & KCK^{-1} \end{pmatrix}$$

eine obere Dreiecksmatrix. Wenn  $W_n = \text{span}(e_1, \dots, e_n)$ , so ist die Fahne von  $A$  gegeben durch  $V_n = M^{-1} \begin{pmatrix} 1 & 0 \\ 0 & K^{-1} \end{pmatrix}$ .

(iii) $\implies$ (i) Ausgehend von der Fahne, wähle  $v_n \in V_n \setminus V_{n-1}$  für  $n = 1, \dots, N$ . Dann ist  $\mathcal{B} = (v_1, \dots, v_N)$  eine geordnete Basis bez. derer  $T_{\mathcal{B}}^{\mathcal{B}}$  eine obere Dreiecksmatrix ist.  $\square$

**6.29 Korollar** Jede Matrix  $A \in \text{Mat}(N \times N, \mathbb{C})$  ist über  $\mathbb{C}$  trigonalisierbar, d.h. es gibt eine invertierbare Matrix  $M$ , so dass  $MAM^{-1}$  eine obere Dreiecksmatrix ist.

**Beweis.** Dies folgt aus Satz 6.28, weil das charakteristische Polynom über  $\mathbb{C}$  nach dem Fundamentalsatz faktorisiert.  $\square$

Als eine Anwendung der Triagonalisierung wird nun ein klassischer Satz bewiesen.

**6.30 Satz** (Satz von Cayley-Hamilton) Sei  $T \in \mathcal{L}(V)$  eine lineare Abbildung auf einem endlich dimensionalen Vektorraum über  $\mathbb{K}$ . Dann gilt für das charakteristische Polynom  $p_T$

$$p_T(T) = 0,$$

wobei  $p_T(T)$  definiert ist wie in Definition 6.21.

**Beweis.** Es reicht, dies für eine darstellende Matrix  $T_B^B$  zu zeigen, da  $p_T(T)_B^B = p_T(T_B^B)$ . Sei  $M \in \text{Mat}(N \times N, \mathbb{C})$  ein Basiswechsel dergestalt, dass  $A = MT_B^B M^{-1}$  eine obere Dreiecksmatrix ist (siehe Korollar 6.29). Es gilt dann

$$p_T(T_B^B) = M^{-1}p_T(A)M = M^{-1}p_A(A)M,$$

Letzteres, weil die charakteristischen Polynome von ähnlichen Matrizen gleich sind. Es reicht also, den Sachverhalt für obere Dreiecksmatrizen zu zeigen. Dies wird durch Induktion über  $N$  getan. Für  $N = 1$  ist es trivial. Für den Schritt von  $N - 1$  nach  $N$  sei  $A$  von der Gestalt

$$A = \begin{pmatrix} \lambda & B \\ 0 & C \end{pmatrix},$$

mit einem Skalar  $\lambda$  und einer oberen Dreiecksmatrix  $C \in \text{Mat}((N - 1) \times (N - 1), \mathbb{K})$ . Nach Induktionsvoraussetzung gilt  $p_C(C) = 0$ . Des Weiteren ist  $p_A(z) = (\lambda - z)p_C(z)$ . Somit folgt für eine geeignete Matrix  $B'$ :

$$p_A(A) = (\lambda \mathbf{1}_N - A)p_C(A) = \begin{pmatrix} 0 & B \\ 0 & C \end{pmatrix} \begin{pmatrix} p_C(\lambda) & B' \\ 0 & p_C(C) \end{pmatrix} = \begin{pmatrix} 0 & B \\ 0 & C \end{pmatrix} \begin{pmatrix} p_C(\lambda) & B' \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

Letzteres durch Ausmultiplizieren. Die zweite Gleichheit benötigt eine gewisse Argumentation, die zur Übung eingefügt werden sollte.  $\square$

**6.31 Definition** Das Spektrum  $\text{Spec}(T) \subset \mathbb{K}$  von  $T \in \mathcal{L}(V)$  ist die Menge aller  $z \in \mathbb{K}$ , für welche  $T - z \mathbf{1}_V$  kein Isomorphismus von  $V$  ist.

In endlich dimensionalen Vektorräumen stimmen Spektrum und die Menge der Eigenwerte überein, wie der folgende Satz zeigt. In unendlich dimensionalen Vektorräumen ist es unbedingt notwendig zu unterscheiden (es gibt kontinuierliches Spektrum, siehe eine Vorlesung über Funktionalanalysis).

**6.32 Satz** Sei  $T \in \mathcal{L}(V)$  auf einem Vektorraum  $V$  über  $\mathbb{K}$  mit endlicher Dimension  $N = \dim(V) < \infty$ . Dann gilt

$$\text{Spec}(T) = \{z \in \mathbb{K} \mid z \text{ Eigenwert von } T\}.$$

**Beweis.** "⊃" Wenn  $z$  Eigenwert ist, so hat  $T - z \mathbf{1}_V$  einen nicht-trivialen Kern und somit ist  $T - z \mathbf{1}_V$  kein Isomorphismus (weil nicht injektiv). "⊂" Wenn  $T - z \mathbf{1}_V$  kein Isomorphismus ist, so ist die lineare Abbildung  $T - z \mathbf{1}_V$  entweder nicht injektiv oder nicht surjektiv. Falls sie nicht injektiv ist, gibt es einen Vektor im Kern, der dann also ein Eigenvektor von  $T$  zum Eigenwert  $z$  ist. Falls  $T - z \mathbf{1}_V$  nicht surjektiv ist, besagt der Rangsatz (Satz 4.8), dass die Abbildung auch nicht injektiv ist, so dass das gleiche Argument zeigt, dass  $z$  Eigenwert ist.  $\square$

## 7 Vektorräume mit Skalarprodukt

**7.1 Definition** Sei  $V$  ein Vektorraum über  $\mathbb{K}$  (wieder ist  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{C}$ ). Dann heißt eine Abbildung  $\langle \cdot | \cdot \rangle : V \times V \rightarrow \mathbb{K}$  Skalarprodukt genau dann, wenn Folgendes für alle  $v, v', w \in V$  und  $\lambda \in \mathbb{K}$  gilt:

$$(i) \quad \langle w | v + \lambda v' \rangle = \langle w | v \rangle + \lambda \langle w | v' \rangle \quad (\text{Linearität im zweiten Argument})$$

$$(ii) \quad \langle w | v \rangle = \overline{\langle v | w \rangle} \quad (\text{Symmetrie})$$

$$(iii) \quad \langle v | v \rangle \geq 0 \quad (\text{Positivität})$$

$$(iv) \quad \langle v | v \rangle = 0 \implies v = 0 \quad (\text{Nicht-Entartung})$$

Des Weiteren wird dann die Norm (auch Länge) eines Vektors definiert als  $\|v\| = \sqrt{\langle v | v \rangle}$ . Ein Vektorraum mit Skalarprodukt wird auch Prä-Hilbertraum genannt.

**7.2 Bemerkung** Ein reeller Vektorraum mit Skalarprodukt wird oft auch als euklidischer Vektorraum bezeichnet, und ein komplexer Vektorraum mit Skalarprodukt als unitärer Vektorraum.  $\diamond$

**7.3 Bemerkung** Die Eigenschaften (i) und (ii) zusammen zeigen, dass ein Skalarprodukt im Fall  $\mathbb{K} = \mathbb{R}$  eine bilineare Abbildung ist, im Sinne von Definition 5.14. Im Fall  $\mathbb{K} = \mathbb{C}$  hingegen ist die Abbildung wegen der komplexen Konjugation nicht linear im ersten Argument, sondern antilinear, d.h. es gilt

$$\langle \lambda v | w \rangle = \bar{\lambda} \langle v | w \rangle .$$

In der Mathematikliteratur wird diese Antilinearität oft auch ins zweite Argument verlegt, in der Physik und insbesondere der Quantenmechanik traditionell nicht. Auch die hier gewählte Notation kommt aus der Physik und ist die sogenannte Dirac'sche BraKet Notation. Ein Bra ist  $\langle v |$  und ein Ket ist  $|w\rangle$ . Zusammen ergibt dies eine *bracket=Klammer*  $\langle v | w \rangle$ . Außerdem ist ein Ket ein Vektor in  $V$ , wohingegen ein Bra ein Vektor aus dem Dualraum  $V^* = \mathcal{L}(V, \mathbb{K})$  ist. In der Mathematikliteratur werden auch andere Notationen für Skalarprodukte verwandt, wie  $\langle v, w \rangle, (v | w) \dots$   $\diamond$

**7.4 Beispiel** Sei  $V = \mathbb{K}^N$ . Das euklidische Skalarprodukt ist dann definiert als

$$\langle v | w \rangle = (\bar{v})^t w = \sum_{n=1}^N \bar{v}_n w_n, \quad v = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix}, \quad w = \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix} .$$

In der Tat gelten alle vier Axiome, z.B. (iii) ist richtig, weil

$$\langle v | v \rangle = \sum_{n=1}^N \bar{v}_n v_n = \sum_{n=1}^N |v_n|^2$$

als Summe nicht-negativer Zahlen nicht-negativ ist. Außerdem kann an dieser Gleichung (iv) abgelesen werden, denn die Summe ist nur genau dann positiv, wenn mindestens ein Summand positiv ist und in diesem Fall ist der Vektor nicht der Nullvektor.  $\diamond$

**7.5 Beispiel** Sei weiter  $V = \mathbb{K}^N$  und  $B \in \text{Mat}(N \times N, \mathbb{K})$ . Wir setzen

$$\langle v | w \rangle_B = (\bar{v})^t B w = \sum_{n=1}^N \bar{v}_n B_{n,m} w_m ,$$

mit  $v, w \in \mathbb{K}^N$  wie oben. Im Fall  $\mathbb{K} = \mathbb{R}$  ist dies genau eine bilineare Abbildung wie definiert in Beispiel 5.15. Auch für  $\mathbb{K} = \mathbb{C}$  ist die Linearität im zweiten Argument offensichtlich gegeben. Die Symmetrie hingegen ist eine Zusatzeigenschaft, die gegeben ist, wenn für  $B$  gilt (was später als Selbstadjungiertheit definiert wird)

$$(\overline{B})^t = B .$$

In der Tat, dann folgt

$$\langle v|w \rangle_B = \sum_{n,m=1}^N \overline{v_n} B_{n,m} w_m = \sum_{n,m=1}^N \overline{v_n} \overline{B_{m,n}} w_m = \overline{\sum_{n=1}^N \overline{w_n} B_{n,m} v_m} = \overline{\langle w|v \rangle_B} .$$

Die Eigenschaften (iii) und (iv) sind dann gleichbedeutend mit der Positivität der Matrix  $B$ , die besagt, dass  $B$  nur positive Eigenwerte hat. Die Positivität von Matrizen wird später im Detail analysiert. Zusammenfassend ist  $\langle \cdot | \cdot \rangle_B$  nur dann ein Skalarprodukt, wenn  $B$  selbstadjungiert und positiv ist. Im Spezialfall  $B = \mathbf{1}_N$  erhält man wieder das euklidische Skalarprodukt.  $\diamond$

**7.6 Beispiel** Sei  $V = \mathbb{R}_N[x]$  der  $N + 1$  dimensionale Vektorraum der Polynome mit reellen Koeffizienten vom Grade höchstens  $N$ . Nun wird das Skalarprodukt definiert mit Hilfe des Integrals als

$$\langle p|q \rangle = \int_0^1 p(x)q(x) dx , \quad p, q \in \mathbb{R}_N[x] .$$

In der Tat gelten alle vier Axiome, wie mit den elementaren Rechenregeln des Integrals überprüft werden kann (Übung). Es ist weiter möglich den Integrationsbereich zu ändern und eine positive Dichte einzuführen (d.h.  $dx$  ersetzen durch  $\rho(x)dx$  mit  $\rho \geq 0$ ) und man erhält weiterhin ein Skalarprodukt, nur eben ein anderes.

Es ist interessant sich dieses Skalarprodukt unter der Koordinatenabbildung  $I_B : \mathbb{R}_N[x] \rightarrow \mathbb{K}^{N+1}$  bez. der Basis  $\mathcal{B} = (1, x, x^2, \dots, x_N)$  der Monome anzusehen. Dann gilt

$$I_{\mathcal{B}}(p) = \begin{pmatrix} p_0 \\ p_1 \\ \vdots \\ p_N \end{pmatrix} , \quad p(x) = \sum_{n=0}^N p_n x^n ,$$

und analog für  $q(x) = \sum_{n=0}^N q_n x^n$ . Nun berechnet man

$$\langle p|q \rangle = \sum_{n,m=0}^N p_n q_m \int_0^1 x^{n+m} dx = \sum_{n,m=0}^N p_n q_m \frac{1}{n+m+1} = I_{\mathcal{B}}(p)^t B I_{\mathcal{B}}(q) ,$$

wobei  $B \in \text{Mat}((N+1) \times (N+1), \mathbb{R})$  gegeben ist durch

$$B = \left( \frac{1}{n+m+1} \right)_{n,m=0,\dots,N} .$$

In der Koordinatendarstellung ist dies also wieder nur ein Spezialfall vom vorangehenden Beispiel, d.h. dieses spezielle  $B$  ist tatsächlich selbstadjungiert und positiv.  $\diamond$

**7.7 Bemerkung** Auf einem endlich dimensionalen Vektorraum  $V$  ist es immer möglich, ein Skalarprodukt zu definieren. Hierzu sei ein beliebige Basis  $\mathcal{B} = (b_1, \dots, b_N)$  gegeben und dann wird solches Skalarprodukt definiert durch

$$\langle b_n|b_m \rangle = \delta_{n,m} ,$$

und anschließend antilineare bzw. lineare Fortsetzung auf  $V \times V$ , d.h.

$$\langle v|w \rangle = \sum_{n=1}^N \overline{v_n} w_n, \quad v = \sum_{n=1}^N v_n b_n, \quad w = \sum_{n=1}^N w_n b_n.$$

Außerdem ist dann  $\mathcal{B}$  eine Orthonormalbasis bez.  $\langle \cdot | \cdot \rangle$  im Sinne folgende Definition. ◇

**7.8 Definition** Sei  $V$  ein Vektorraum mit Skalarprodukt. Dann definieren wir

- (i)  $v, w \in V$  orthogonal (Kurzschreibweise  $v \perp w$ )  $\iff \langle v|w \rangle = 0$ .
- (ii)  $v \in V$  normiert  $\iff \|v\| = \sqrt{\langle v|v \rangle} = 1$ .
- (iii)  $(u_i)_{i \in I}$  orthogonale Familie  $\iff u_i \perp u_j \forall i, j \in I, i \neq j$ .
- (iv)  $(u_i)_{i \in I}$  orthonormierte Familie  $\iff (u_i)_{i \in I}$  orthogonale Familie und  $\|u_i\| = 1 \forall i \in I$   
 $\iff \langle u_i|u_j \rangle = \delta_{i,j} \forall i, j \in I$ .
- (v) Eine Orthonormalbasis ist eine Basis, die eine orthonormierte Familie ist.

**7.9 Beispiel** Wenn der  $\mathbb{K}^N$  mit dem euklidischen Skalarprodukt versehen wird, dann ist die Standardbasis eine Orthonormalbasis (dies sei zur Übung überprüft). ◇

**7.10 Satz** Seien  $(u_k)_{k=1, \dots, K}$  orthonormiert. Dann gilt für alle  $v \in V$ :

$$\|v\|^2 = \sum_{k=1}^K |\langle v|u_k \rangle|^2 + \|v - \sum_{k=1}^K \langle u_k|v \rangle u_k\|^2 \quad (\text{Pythagoras}),$$

und

$$\|v\|^2 \geq \sum_{k=1}^K |\langle v|u_k \rangle|^2 \quad (\text{Bessels Ungleichung}).$$

**Beweis:** Setze  $w = \sum_{k=1}^K \langle u_k|v \rangle u_k$ . Dann gilt  $(v - w) \perp w$ , da

$$\begin{aligned} \langle v - w|w \rangle &= \sum_{k=1}^K \langle u_k|v \rangle \langle v - w|u_k \rangle \\ &= \sum_{k=1}^K \langle u_k|v \rangle \left( \langle v|u_k \rangle - \sum_{l=1}^K \overline{\langle u_l|v \rangle} \underbrace{\langle u_l|u_k \rangle}_{\delta_{l,k}} \right) = 0. \end{aligned}$$

Geometrisch kann  $w$  als die orthogonale Projektion von  $v$  auf den von  $u_1, \dots, u_K$  aufgespannten Unterraum verstanden werden. Somit

$$\|v\|^2 = \langle v|v \rangle = \langle v - w + w|v - w + w \rangle = \langle v - w|v - w \rangle + \langle w|w \rangle,$$

was den Beweis beendet. □

**7.11 Korollar** (Cauchy-Schwarz Ungleichung) Sei  $V$  ein Vektorraum mit Skalarprodukt und seien  $v, w \in V$ . Dann

$$|\langle v|w \rangle| \leq \|v\| \|w\|.$$

**Beweis:** Der Fall mit  $w = 0$  ist trivial. Sonst wende man die Bessel-Ungleichung auf die einelementige orthonormale Menge  $\frac{w}{\|w\|}$  an:

$$\|v\|^2 \geq |\langle v | \frac{w}{\|w\|} \rangle|^2 = \frac{|\langle v | w \rangle|^2}{\|w\|^2},$$

was schon die Ungleichung ist. □

**7.12 Bemerkung** Die Cauchy-Schwarz Ungleichung erlaubt es zwischen je zwei Vektoren  $v, w \in V$  einen Winkel  $\sphericalangle(v, w) \in [0, \frac{\pi}{2}]$  zu definieren durch

$$\cos(\sphericalangle(v, w)) = \frac{|\langle v | w \rangle|}{\|v\| \|w\|}.$$

In der Tat, die rechte Seite liegt in  $[0, 1]$  und  $\cos : [0, \frac{\pi}{2}] \rightarrow [0, 1]$  ist bijektiv. ◇

**7.13 Satz** Sei  $V$  ein Vektorraum mit Skalarprodukt. Dann erfüllt die Norm  $\|v\| = \sqrt{\langle v | v \rangle}$ :

- (i)  $\|\lambda v\| = |\lambda| \|v\|$  (Homogenität)
- (ii)  $\|v + w\| \leq \|v\| + \|w\|$  (Dreiecksungleichung)
- (iii)  $\|v\| = 0 \implies v = 0$  (Nicht-Entartung)

Mit diesen Eigenschaften wird  $(V, \|\cdot\|)$  zu einem sogenannten normierten Vektorraum.

**Beweis:**  $\|\lambda v\| = |\lambda| \|v\|$  ist klar, ebenso wie die Positivität und die Nicht-Entartung. Nun zur Dreiecksungleichung:

$$\begin{aligned} \|v + w\|^2 &= \langle v | v \rangle + \langle v | w \rangle + \langle w | v \rangle + \langle w | w \rangle \\ &= \|v\|^2 + 2 \Re \langle v | w \rangle + \|w\|^2 \\ &\leq \|v\|^2 + 2 \|v\| \|w\| + \|w\|^2 = (\|v\| + \|w\|)^2, \end{aligned}$$

wobei wir die Cauchy-Schwarz-Ungleichung verwandt haben. □

**7.14 Korollar** Sei  $V$  ein Vektorraum mit Skalarprodukt und  $v \in V$ . Dann

$$\|v\| = \sup_{\|w\|=1} |\langle w | v \rangle|.$$

**Beweis:** Die Cauchy-Schwarz Ungleichung impliziert  $|\langle w | v \rangle| \leq \|w\| \|v\| = \|v\|$  und somit ist das Supremum auch kleiner oder gleich  $\|v\|$ . Andererseits führt die Wahl  $w = v/\|v\|$  zur Gleichheit. □

Nun zu einer weiteren geometrischen Information.

**7.15 Satz** Sei  $(V, \langle \cdot | \cdot \rangle)$  ein Vektorraum mit Skalarprodukt. Dann gilt

- (i)  $\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2)$  (Parallelogrammregel)
- (ii) Das Skalarprodukt kann durch die Norm  $\|\cdot\|$  berechnet werden mit der Polarisationsidentität:

$$\langle v | w \rangle = \frac{1}{4} [(\|v + w\|^2 - \|v - w\|^2) - i(\|v + iw\|^2 - \|v - iw\|^2)], \quad \text{falls } \mathbb{K} = \mathbb{C},$$

und

$$\langle v | w \rangle = \frac{1}{4} [\|v + w\|^2 - \|v - w\|^2], \quad \text{falls } \mathbb{K} = \mathbb{R}.$$

**Beweis:** (i) ist eine rein algebraische Rechnung:

$$\begin{aligned} \|v+w\|^2 + \|v-w\|^2 &= \langle v+w|v+w\rangle + \langle v-w|v-w\rangle \\ &= \langle v|v\rangle + \langle v|w\rangle + \langle w|v\rangle + \langle w|w\rangle + \langle v|v\rangle - \langle v|w\rangle - \langle w|v\rangle + \langle w|w\rangle \\ &= 2\langle v|v\rangle + 2\langle w|w\rangle, \end{aligned}$$

und (ii) verifiziert man analog.  $\square$

**7.16 Bemerkung** Nach einem etwas tiefer liegenden Satz von John von Neumann wird die Norm eines normierten Vektorraumes genau dann durch ein Skalarprodukt induziert, wenn die Norm die Parallelogrammregel erfüllt. Das Skalarprodukt ist dann durch Satz 7.15 (ii) gegeben.  $\diamond$

Weiter oben wurden orthonormale Familien schon im Beweis verwandt. Der folgende Sachverhalt zeigt, dass es sehr einfach ist, Vektoren nach einer Orthonormalbasis zu entwickeln.

**7.17 Satz** Sei  $V$  ein endlich-dimensionaler Vektorraum mit Skalarprodukt und  $\mathcal{B} = (u_1, \dots, u_N)$  eine geordnete Orthonormalbasis. Dann gilt für die Koordinatenabbildung  $I_{\mathcal{B}} : V \rightarrow \mathbb{K}^N$

$$I_{\mathcal{B}}(v) = \begin{pmatrix} \langle u_1|v\rangle \\ \vdots \\ \langle u_N|v\rangle \end{pmatrix}.$$

Außerdem

$$\langle v|w\rangle = \langle I_{\mathcal{B}}(v)|I_{\mathcal{B}}(w)\rangle, \quad v, w \in V, \quad (7.1)$$

wobei auf der rechten Seite das euklidische Skalarprodukt im  $\mathbb{K}^N$  verwandt wurde.

**Beweis:** Es sei an die Definition der Koordinatenabbildung erinnert:

$$I_{\mathcal{B}}(v) = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix}, \quad \text{wenn } v = \sum_{k=1}^N v_k u_k.$$

Wenn nun das Skalarprodukt letzterer Gleichung mit  $u_n$  gebildet wird, folgt

$$\langle u_n|v\rangle = \sum_{k=1}^N v_k \langle u_n|u_k\rangle = v_n,$$

und somit die erste Behauptung. Auch die zweite folgt aus einer Rechnung:

$$\langle v|w\rangle = \left\langle \sum_{k=1}^N \langle u_k|v\rangle u_k \mid \sum_{n=1}^N \langle u_n|w\rangle u_n \right\rangle = \sum_{k,n=1}^N \overline{\langle u_k|v\rangle} \langle u_n|w\rangle \delta_{n,k} = \sum_{n=1}^N \langle v|u_n\rangle \langle u_n|w\rangle, \quad (7.2)$$

und Letzteres ist genau  $\langle I_{\mathcal{B}}(v)|I_{\mathcal{B}}(w)\rangle$ .  $\square$

**7.18 Bemerkung** Somit müssen bei einer Orthonormalbasis lediglich Skalarprodukte berechnet werden und kein Gleichungssystem gelöst werden, wenn die Entwicklung eines Vektors nach einer Basis bestimmt werden soll. Es ist möglich, den Sachverhalt aus Satz 7.17 auch mit der Dirac'schen BraKet Notation zu

schreiben. Seien dazu  $|n\rangle = e_n$  die Standardbasis Vektoren in  $\mathbb{K}^N$ , die ja auch eine Orthonormalbasis bez. des euklidischen Skalarproduktes bilden. Dann gilt

$$I_{\mathcal{B}} = \sum_{n=1}^N |n\rangle\langle u_n|. \quad (7.3)$$

Dies ist so zu verstehen, dass die Wirkung auf eine Ket-Vektor  $v = |v\rangle \in V$  gegeben ist durch

$$I_{\mathcal{B}} v = I_{\mathcal{B}}|v\rangle = \sum_{n=1}^N |n\rangle\langle u_n|v\rangle = \sum_{n=1}^N \langle u_n|v\rangle e_n,$$

wobei also die BraKet's als Koeffizienten auftreten. Beachte, dass wir hier zwischen Dirac Notation und konventioneller Notation hin und hergesprungen sind. Ähnlich wie (7.3) sind auch folgende Identitäten zu überprüfen:

$$\mathbf{1}_N = \sum_{n=1}^N |n\rangle\langle n| \in \text{Mat}(N \times N, \mathbb{K}), \quad \mathbf{1}_V = \sum_{n=1}^N |u_n\rangle\langle u_n| \in \mathcal{L}(V).$$

In der Tat gilt z.B.

$$\left( \sum_{n=1}^N |u_n\rangle\langle u_n| \right) v = \sum_{n=1}^N |u_n\rangle\langle u_n|v\rangle = \sum_{n=1}^N \langle u_n|v\rangle u_n = v,$$

Letzteres weil ja gerade die Entwicklung nach der Basis  $\mathcal{B}$  vorliegt. Die Identität folgt auch aus (7.2).  $\diamond$

Nachdem nun einige erste Anwendungen von Orthonormalbasen aufgezeigt wurden, soll das Standardverfahren zur Erzeugung von solchen Basen vorgestellt werden.

**7.19 Satz (Gram-Schmidt Verfahren)** Sei  $V$  ein Vektorraum mit Skalarprodukt. Wenn  $v_1, \dots, v_N$  eine Menge linear unabhängiger Vektoren in  $V$  ist, so gibt es eine orthonormierte Familie  $u_1, \dots, u_N$  mit

$$\text{span}(\{v_1, \dots, v_n\}) = \text{span}(\{u_1, \dots, u_n\}), \quad n = 1, \dots, N.$$

Die orthonormierte Familie kann iterativ konstruiert werden durch

$$u_n = \frac{v_n - \sum_{k=1}^{n-1} \langle u_k|v_n\rangle u_k}{\|v_n - \sum_{k=1}^{n-1} \langle u_k|v_n\rangle u_k\|}.$$

**Beweis:** Setze  $w_n = \sum_{k=1}^{n-1} \langle u_k|v_n\rangle u_k$ . Dies ist die Projektion von  $v_n$  auf den von  $u_1, \dots, u_{n-1}$  aufgespannten Unterraum, d.h.  $(v_n - w_n) \perp u_k$  für alle  $k = 1, \dots, n-1$ , denn nach der gleichen Rechnung wie im Beweis von Satz 7.10:

$$\langle u_k|v_n - w_n\rangle = \langle u_k|v_n\rangle - \sum_{l=1}^{n-1} \langle u_l|v_n\rangle \langle u_k|u_l\rangle = \langle u_k|v_n\rangle - \sum_{l=1}^{n-1} \langle u_l|v_n\rangle \delta_{k,l} = 0.$$

Also bilden  $u_1, \dots, u_{n-1}, v_n - w_n$  eine orthogonale Familie. Normieren des letzten Vektors ergibt  $u_n$  und somit eine orthonormierte Familie.  $\square$

Zusammen mit dem Basiserweiterungssatz erhält man:



**7.20 Korollar** Jede orthonormierte Familie kann zu einer Orthonormalbasis erweitert werden.

**7.21 Beispiel** Seien  $V = \mathbb{R}^3$  und

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Diese Vektoren bilden eine Basis des  $\mathbb{R}^3$ , da sie linear unabhängig sind. Das Gram-Schmidt Verfahren wird durchgeführt, um eine Orthonormalbasis  $u_1, u_2, u_3$  zu erzeugen. Zunächst wird also  $v_1$  normiert:

$$u_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Nun wird zunächst berechnet

$$w_2 = v_2 - \langle v_2 | u_1 \rangle u_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix},$$

und  $u_2$  ergibt sich wieder durch Normierung:

$$u_2 = \frac{w_2}{\|w_2\|} = \frac{1}{\sqrt{\frac{1}{4} + \frac{1}{4} + 1}} \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix} = \frac{2}{\sqrt{6}} \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix}.$$

Genauso wird für  $u_3$  verfahren. Zunächst

$$\begin{aligned} w_3 &= v_3 - \langle v_3 | u_1 \rangle u_1 - \langle v_3 | u_2 \rangle u_2 \\ &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - 0 - \frac{2}{\sqrt{6}} \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \middle| \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix} \right\rangle \frac{2}{\sqrt{6}} \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \frac{2}{6} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix} \end{aligned}$$

und wiederum durch Normieren

$$u_3 = \frac{w_3}{\|w_3\|} = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$

Somit ist die Orthonormalbasis berechnet. ◇

**7.22 Bemerkung** Mit Hilfe des Gram-Schmidt Verfahrens werden sämtliche Familien orthogonaler Polynome erzeugt (Legendre, Hermite, Laguerre, etc.). ◇

**7.23 Satz** (Endlich dimensionales Riesz Lemma) Seien  $V$  ein endlich dimensionaler Vektorraum mit Skalarprodukt und  $T \in \mathcal{L}(V, \mathbb{K})$ . Dann gibt es einen eindeutig bestimmten Vektor  $w_T \in V$  (der  $T$  spezifiziert), so dass

$$T(v) = \langle w_T | v \rangle, \quad v \in V.$$

**Beweis:** Als lineare Abbildung ist  $T$  durch seine Werte auf einer Orthonormalbasis  $u_1, \dots, u_N$  festgelegt (siehe Satz 4.26). Wähle

$$w_T = \sum_{n=1}^N \overline{T(u_n)} u_n .$$

Einsetzen zeigt dann  $T(u_n) = \langle w_T | u_n \rangle$  für  $n = 1, \dots, N$ . Nun sind beide Seiten  $T(v) = \langle w_T | v \rangle$  linear in  $v$ . Da sie auf einer Basis übereinstimmen, sind sie als lineare Abbildungen gleich.  $\square$

**7.24 Definition** Sei  $V$  ein Vektorraum über  $\mathbb{K}$  (nicht notwendigerweise versehen mit einem Skalarprodukt). Eine lineare Abbildung  $T : V \rightarrow \mathbb{K}$  heißt auch lineares Funktional und der  $\mathbb{K}$ -Vektorraum  $\mathcal{L}(V, \mathbb{K})$  heißt der Dualraum von  $V$ , der meist bezeichnet wird mit  $V^* = \mathcal{L}(V, \mathbb{K})$ .

**7.25 Bemerkung** Die Abbildung  $T \in V^* \mapsto w_T \in V$ , deren Existenz im Riesz Lemma bewiesen wird, ist ein Isomorphismus zwischen  $V^*$  und  $V$ , d.h. eine bijektive lineare Abbildung. Auch wenn ein endlich dimensionaler Vektorraum  $V$  kein Skalarprodukt hat, gibt es einen Isomorphismus zwischen  $V$  und  $V^*$  (tatsächlich gibt es unendlich viele). In der Tat, sei  $(b_1, \dots, b_N)$  eine Basis von  $V$ . Dann definiere  $b_n^* \in V^*$  durch seine Werte auf der Basis von  $V$  durch

$$b_n^*(b_j) = \delta_{n,j} .$$

Nun sei als Übung überprüft, dass  $b_1^*, \dots, b_N^*$  eine Basis bilden von  $V^*$  (insbesondere gilt es also zu zeigen, dass jedes  $T \in V^*$  Linearkombination von  $b_1^*, \dots, b_N^*$  ist). Also haben die Basen von  $V$  und  $V^*$  gleich viele Elemente und somit sind  $V$  und  $V^*$  als  $\mathbb{K}$ -Vektorräume gleicher Dimension isomorph.  $\diamond$

**7.26 Satz** Seien  $V, W, U$  endlich dimensionale Vektorräume mit Skalarprodukt und  $T \in \mathcal{L}(V, W)$ . Dann gibt es eine eindeutig bestimmte lineare Abbildung  $T^* \in \mathcal{L}(W, V)$ , so dass

$$\langle T^* w | v \rangle = \langle w | T v \rangle , \quad v \in V , \quad w \in W .$$

Die lineare Abbildung  $T^*$  heißt die adjungierte Abbildung zu  $T$ . Wenn zudem  $S \in \mathcal{L}(W, U)$ , dann

$$(ST)^* = T^* S^* .$$

**Beweis:** Es wird zunächst die Existenz gezeigt. Die Abbildung  $v \in V \mapsto \langle w | T v \rangle$  ist ein lineares Funktional, also existiert nach dem Lemma von Riesz ein  $v_w \in V$ , so dass

$$\langle w | T v \rangle = \langle v_w | v \rangle .$$

Nun setze  $T^* w = v_w$ . Zunächst zeigen wir, dass das so definierte  $T^*$  linear ist. Für die Linearkombination  $w + \lambda w'$  gilt

$$\begin{aligned} \langle T^*(w + \lambda w') | v \rangle &= \langle v_{w+\lambda w'} | v \rangle = \langle w + \lambda w' | T v \rangle = \langle w | T v \rangle + \lambda \langle w' | T v \rangle \\ &= \langle v_w | v \rangle + \lambda \langle v_{w'} | v \rangle = \langle T^* w | v \rangle + \lambda \langle T^* w' | v \rangle . \end{aligned}$$

Also gilt

$$\langle T^*(w + \lambda w') - T^* w - \lambda T^* w' | v \rangle = 0 , \quad \forall v \in V .$$

Hieraus folgt nach unten stehendem Lemma die Linearität von  $T^*$ . Die Eindeutigkeit von  $T^*$  sowie die letzte Behauptung seien als Übung überprüft.  $\square$

**7.27 Lemma** Sei  $V$  ein Vektorraum mit Skalarprodukt und  $v \in V$ . Falls  $\langle w | v \rangle = 0$  für alle  $w \in V$ , so gilt  $v = 0$ .

**Beweis:** In der Tat gilt insbesondere  $\langle v|v \rangle = 0$  und somit  $v = 0$  nach der Nichtentartung des Skalarproduktes.  $\square$

**7.28 Definition** Zu  $A \in \text{Mat}(N \times M, \mathbb{K})$  ist die adjungierte Matrix  $A^* \in \text{Mat}(M \times N, \mathbb{K})$  definiert als  $A^* = (\overline{A})^t$ , d.h.

$$(A^*)_{n,m} = \overline{A_{m,n}}.$$

**7.29 Satz** Sei  $T \in \mathcal{L}(V, W)$  eine lineare Abbildung zwischen zwei Vektorräumen  $V, W$  mit Skalarprodukt. Seien  $\mathcal{B}$  und  $\mathcal{C}$  Orthonormalbasen von  $V$  und  $W$ . Dann gilt

$$(T^*)_{\mathcal{B}}^{\mathcal{C}} = (T_{\mathcal{C}}^{\mathcal{B}})^*,$$

wobei links die adjungierte Abbildung und rechts die adjungierte Matrix auftaucht.

Insbesondere, wenn  $V = \mathbb{K}^N$  und  $W = \mathbb{K}^M$  und  $\mathcal{B}$  und  $\mathcal{C}$  die Standardbasen sind, dann ist die adjungierte Abbildung zu einer linearen Abbildung  $A \in \text{Mat}(N \times M, \mathbb{K})$  gegeben durch die Matrixmultiplikation mit der adjungierten Matrix.

**Beweis:** Es wird (7.1) verwandt, d.h.

$$\langle v|v' \rangle = \langle I_{\mathcal{B}}(v)|I_{\mathcal{B}}(v') \rangle, \quad v, v' \in V,$$

und analog für  $W$ . Außerdem sei daran erinnert, dass  $T_{\mathcal{C}}^{\mathcal{B}} = I_{\mathcal{C}}T(I_{\mathcal{B}})^{-1}$ . Also, unter Verwendung der adjungierten Matrix,

$$\begin{aligned} \langle w|Tv \rangle &= \langle I_{\mathcal{C}}(w)|I_{\mathcal{C}}(Tv) \rangle = \langle I_{\mathcal{C}}(w)|T_{\mathcal{C}}^{\mathcal{B}}I_{\mathcal{B}}(v) \rangle = \sum_{n=1}^{\dim(W)} \sum_{m=1}^{\dim(V)} \overline{w_n} (T_{\mathcal{C}}^{\mathcal{B}})_{n,m} v_m \\ &= \sum_{n=1}^{\dim(W)} \sum_{m=1}^{\dim(V)} \overline{((T_{\mathcal{C}}^{\mathcal{B}})^*)_{m,n} w_n} v_m = \langle (T_{\mathcal{C}}^{\mathcal{B}})^* I_{\mathcal{C}}(w)|I_{\mathcal{B}}(v) \rangle = \langle (I_{\mathcal{B}}^{-1}(T_{\mathcal{C}}^{\mathcal{B}})^* I_{\mathcal{C}})w|v \rangle. \end{aligned}$$

Da die adjungierte Abbildung nach Satz 7.26 eindeutig ist, folgt  $I_{\mathcal{B}}^{-1}(T_{\mathcal{C}}^{\mathcal{B}})^* I_{\mathcal{C}} = T^*$ , und somit gilt auch  $(T_{\mathcal{C}}^{\mathcal{B}})^* = I_{\mathcal{B}}T^*I_{\mathcal{C}}^{-1} = (T^*)_{\mathcal{B}}^{\mathcal{C}}$ .  $\square$

**7.30 Definition** Seien  $V$  und  $W$  endlich dimensionale Vektorräume über  $\mathbb{K}$  mit Skalarprodukt.

- (i) Eine lineare Abbildung  $U \in \mathcal{L}(V, W)$  heißt unitär genau dann, wenn  $U^*U = \mathbf{1}_V$  und  $UU^* = \mathbf{1}_W$ .
- (ii) Falls  $\mathbb{K} = \mathbb{R}$ , heißen unitäre Abbildungen auch orthogonal.
- (iii) Eine lineare Abbildung  $T \in \mathcal{L}(V)$  heißt normal genau dann, wenn  $T^*T = TT^*$ .
- (iv) Eine lineare Abbildung  $T \in \mathcal{L}(V)$  heißt selbstadjungiert oder hermitisch genau dann, wenn  $T^* = T$ .

Bei Matrizen verwenden all diese Begriffsbildungen das euklidische Skalarprodukt, so dass die adjungierte Matrix wie in Definition 7.28 gegeben ist.

**7.31 Bemerkungen** 1. Jede selbstadjungierte Abbildung ist normal.

2. Jede unitäre Abbildung  $U \in \mathcal{L}(V) = \mathcal{L}(V, V)$  ist normal.

3. Für jede unitäre Abbildung  $U \in \mathcal{L}(V, W)$  gilt  $U^* = U^{-1}$ . Insbesondere folgt aus  $U^*U = \mathbf{1}_V$  auch  $UU^* = \mathbf{1}_W$ , und umgekehrt.

4. Jede unitäre Abbildung  $U \in \mathcal{L}(V, W)$  erhält das Skalarprodukt, d.h.

$$\langle Uv|Uw \rangle = \langle v|w \rangle, \quad \forall v, w \in V.$$

Umgekehrt, wenn eine Abbildung  $U \in \mathcal{L}(V, W)$  alle Skalarprodukte erhält, so ist sie unitär.

5. Wenn  $\mathcal{B} = (u_1, \dots, u_N)$  eine Orthonormalbasis von  $V$  ist, so ist  $U = (u_1, \dots, u_N) \in \mathcal{L}(\mathbb{K}^N, V)$  ein unitäre Abbildung, weil

$$U^*U = (\langle e_n|U^*Ue_m \rangle)_{n,m=1,\dots,N} = (\langle Ue_n|Ue_m \rangle)_{n,m=1,\dots,N} = (\langle u_n|u_m \rangle)_{n,m=1,\dots,N} = \mathbf{1}.$$

Umgekehrt, wenn  $U \in \mathcal{L}(\mathbb{K}^N, V)$ , dann bilden  $Ue_1, \dots, Ue_N$  eine Orthonormalbasis von  $V$ .

6. Die unitären Abbildungen auf  $V$  bilden eine Gruppe, wenn als Gruppenoperation die Hintereinanderausführung gewählt wird, denn wenn  $U, U'$  unitär sind, so auch  $UU'$  weil  $(UU')^*(UU') = (U')^*U^*UU' = (U')^*U' = \mathbf{1}$ . Diese Gruppe heißt die unitäre Gruppe von  $V$  und wird mit  $\mathcal{U}(V) = \{U \in \mathcal{L}(V) \mid U \text{ unitär}\}$  bezeichnet.

7. Wenn  $V = \mathbb{C}^N$  versehen mit dem Standardskalarprodukt, dann wird die unitäre Gruppe meist mit  $U(N)$  bezeichnet und heißt die  $N$ -dimensionale unitäre Gruppe. Wenn  $V = \mathbb{R}^N$  versehen mit dem Standardskalarprodukt, dann wird die unitäre Gruppe meist mit  $O(N)$  bezeichnet und heißt dann auch die orthogonale Gruppe der Dimension  $N$ .  $\diamond$

**7.32 Satz** (Unitäre Trigonalisierung nach Schur) Sei  $A \in \text{Mat}(N \times N, \mathbb{C})$ . Dann existiert eine unitäre Matrix  $U \in \text{Mat}(N \times N, \mathbb{C})$ , so dass  $U^*AU$  eine obere Dreiecksmatrix ist.

**Beweis:** Der Beweis ist ganz analog zu dem Beweis der Implikation (ii) $\implies$ (iii) in Satz 6.28, nur dass noch die Unitarität als Zusatzeigenschaft verwandt wird. Dennoch führen wir den Beweis im Detail. Wieder wird eine Induktion über  $N$  gemacht und der Fall  $N = 1$  ist offensichtlich, so dass sogleich der Schritt  $N - 1 \rightarrow N$  betrachtet wird. Sei also  $A \in \text{Mat}(N \times N, \mathbb{C})$ . Da der Körper  $\mathbb{C}$  ist, gibt es mindestens einen Eigenwert  $z_1$  mit Eigenvektor  $u_1$ , den wir als normiert annehmen dürfen. Seien  $u_2, \dots, u_N \in \mathbb{C}^N$ , so dass  $u_1, \dots, u_N$  eine Orthonormalbasis sind (siehe Korollar 7.20). Dann gilt mit  $V = (u_1, \dots, u_N)$

$$V^*AV = (u_1, \dots, u_N)^*(Au_1, \dots, Au_N) = (u_1, \dots, u_N)^*(z_1u_1, Au_2, \dots, Au_N) = \begin{pmatrix} z_1 & B \\ 0 & C \end{pmatrix},$$

wobei  $C \in \text{Mat}((N - 1) \times (N - 1), \mathbb{C})$ . Für diese Abbildung auf dem  $\mathbb{C}^{N-1}$  gibt es nach Induktionsvoraussetzung eine unitäres  $W \in \text{Mat}((N - 1) \times (N - 1), \mathbb{C})$ , so dass  $W^*CW$  eine obere Dreiecksmatrix ist. Also ist

$$\begin{pmatrix} 1 & 0 \\ 0 & W^* \end{pmatrix} V^*AV \begin{pmatrix} 1 & 0 \\ 0 & W \end{pmatrix} = \begin{pmatrix} z_1 & BW \\ 0 & W^*CW \end{pmatrix}$$

eine obere Dreiecksmatrix. Nun ist aber auch  $U = V \begin{pmatrix} 1 & 0 \\ 0 & W \end{pmatrix}$  unitär als Produkt von unitären.  $\square$

**7.33 Satz** (Spektralsatz für normale Matrizen) Sei  $A \in \text{Mat}(N \times N, \mathbb{C})$ . Dann gilt

$$A \text{ normal} \iff A \text{ unitär diagonalisierbar},$$

wobei Letzteres bedeutet, dass es Eigenwerte  $z'_1, \dots, z'_N \in \mathbb{C}$  gibt (nicht notwendigerweise paarweise verschieden) und eine unitäre Matrix  $U$ , so dass

$$U^*AU = D, \quad D = \begin{pmatrix} z'_1 & & \\ & \ddots & \\ & & z'_N \end{pmatrix}.$$

Für normales  $A$  gelten zudem die Äquivalenzen:

$$\begin{aligned} A \text{ selbstadjungiert} &\iff z'_1, \dots, z'_N \in \mathbb{R}, \\ A \text{ unitär} &\iff 1 = |z'_1| = \dots = |z'_N|. \end{aligned}$$

**Beweis:** "←" Weil Diagonalmatrizen vertauschen gilt dann

$$A^*A = UD^*U^*UDU^* = UD^*DU^* = UDD^*U^* = UDU^*UD^*U^* = AA^*.$$

"⇒" Nach Satz 7.32 gibt es eine unitäre Matrix  $U$ , so dass  $B = U^*AU$  eine obere Dreiecksmatrix ist. Es folgt dann aus der Voraussetzung

$$B^*B = U^*A^*UU^*AU = U^*A^*AU = U^*AA^*U = U^*AUU^*A^*U = BB^*,$$

d.h.  $B$  ist auch normal. Nun sind aber normale obere Dreiecksmatrizen diagonal. Dies wird wieder durch Induktion über  $N$  gezeigt. Der Anfang  $N = 1$  ist klar. Für den Schritt von  $N - 1 \rightarrow N$  betrachten wir

$$|B_{1,1}|^2 = (B^*B)_{1,1} = (BB^*)_{1,1} = |B_{1,1}|^2 + \sum_{n=2}^N |B_{1,n}|^2,$$

so dass  $B_{1,n} = 0$  für  $n = 2, \dots, N$ , d.h. in der ersten Zeile von  $B$  kann nur der Diagonaleintrag  $B_{1,1}$  ungleich Null sein. Somit

$$B = \begin{pmatrix} B_{1,1} & 0 \\ 0 & C \end{pmatrix},$$

wobei  $C$  eine normal obere Dreiecksmatrix der Größe  $N-1$  ist, also nach Induktionsvoraussetzung diagonal ist. Somit ist der Beweis beendet.

Nun zum Zusatz. Sei also  $A$  selbstadjungiert und  $Av = zv$  mit  $\|v\| = 1$ . Dann gilt

$$z = z \langle v|v \rangle = \langle v|zv \rangle = \langle v|Av \rangle = \langle A^*v|v \rangle = \langle Av|v \rangle = \langle zv|v \rangle = \bar{z}.$$

Umgekehrt, wenn  $z'_1, \dots, z'_N \in \mathbb{R}$ , dann gilt  $A^* = (UDU^*)^* = UD^*U^* = UDU^* = A$ . Ähnlich wird bei unitärem  $A$  verfahren.  $\square$

### 7.34 Beispiel Sei

$$A = \begin{pmatrix} a & b \\ b & a \end{pmatrix},$$

wobei  $a, b \in \mathbb{C}$ . Dann zeigt eine kurze Rechnung, dass  $AA^* = A^*A$  gilt und somit  $A$  normal ist. Also kann  $A$  unitär diagonalisiert werden und dies soll nun angegeben werden. Das charakteristische Polynom ist  $p_A(z) = (a-z)^2 - b^2$  und somit sind  $z_1 = a+b$  und  $z_2 = a-b$  die Eigenwerte. Zugehörige normalisierte Eigenvektoren sind

$$v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Tatsächlich sind diese Vektoren auch orthogonal und somit ist

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

auch unitär. Dann ist  $U^{-1}AU = U^*AU = D$  diagonal. Die Eigenvektoren sind nur bis auf eine Phase eindeutig. Insbesondere kann anstelle von  $U$  auch verwandt werden

$$U' = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{i\theta_1} & e^{i\theta_2} \\ e^{i\theta_1} & -e^{i\theta_2} \end{pmatrix}, \quad \theta_1, \theta_2 \in \mathbb{R}.$$

es gilt dann immer noch  $(U')^*AU' = D$ . Falls  $b = 0$  (dann ist  $A = 1$ ) kann  $U$  sogar also beliebiges Element in  $U(2)$  gewählt werden. Diese Uneindeutigkeiten bei der Wahl von  $U$  (Wahl von Orthonormalbasen in evtl. entarteten Eigenräumen) treten immer bei der unitären Diagonalisierung normaler Matrizen auf.  $\diamond$

**7.35 Satz** (Spektralsatz für reelle symmetrische Matrizen) Sei  $A = A^t \in \text{Mat}(N \times N, \mathbb{R})$  symmetrisch. Dann gibt eine reelle unitäre Matrix  $O$ , auch genannt eine orthogonale Matrix, so dass

$$O^t A O = D, \quad D = \begin{pmatrix} z'_1 & & \\ & \ddots & \\ & & z'_N \end{pmatrix},$$

wobei  $z'_1, \dots, z'_N \in \mathbb{R}$  die Eigenwerte von  $A$  sind.

**Beweis:** Zunächst wird  $A$  als komplexe Matrix aufgefaßt. Diese ist selbstadjungiert, denn  $A^* = (\bar{A})^t = A^t = A$ . Also existiert nach Satz 7.33 eine unitäre (komplexe) Matrix  $U$ , so dass  $U^*AU = D$ . Außerdem sind die Spaltenvektoren  $u_n$  von  $U = (u_1, \dots, u_N)$  Eigenvektoren von  $A$ . Zu einem Eigenwert  $z_k$  sei  $E_A(z_k)$  der zugehörige Eigenraum, der also von einigen der Vektoren  $u_n$  aufgespannt wird. Für jeden Vektor  $v \in E_A(z_k)$  gilt nun

$$Av = z_k v \quad \iff \quad A\bar{v} = z_k \bar{v}.$$

Also ist  $E_A(z_k)$  invariant unter der komplexen Konjugation  $v \mapsto \bar{v}$ . Nach dem unten stehenden Lemma existiert eine reelle Orthonormalbasis  $v_1, \dots, v_{\alpha(z_k)}$  von  $E_A(z_k)$ . Wenn diese Orthonormalbasen für alle Eigenwerte vereinigt werden, entsteht eine Orthonormalbasis  $(v_1, \dots, v_N)$  von  $\mathbb{R}^N$ , die also eine orthogonale Matrix  $O = (v_1, \dots, v_N)$  mit den gewünschten Eigenschaften liefern.  $\square$

**7.36 Lemma** Sei  $V \subset \mathbb{C}^N$  ein Unterraum, der invariant unter komplexer Konjugation ist, d.h.  $\bar{V} = V$ . Dann existiert eine reelle Orthonormalbasis von  $V$ .

**Beweis:** Die Basis wird iterativ konstruiert. Sei  $V' \subset V$  ein Unterraum, der schon eine reelle Orthonormalbasis  $v_1, \dots, v_n$  besitzt. Wähle einen Einheitsvektor  $u \in V$ , der orthogonal auf  $V'$  steht. Dann sind  $\Re(u) = \frac{1}{2}(u + \bar{u})$  und  $\Im(u) = \frac{1}{2i}(u - \bar{u})$  beide in  $V$  und sind beide orthogonal auf  $V'$ , weil

$$\langle v_k | \bar{u} \rangle = \overline{\langle v_k | u \rangle} = 0.$$

Nun ist sicher einer der beiden Vektoren  $\Re(u)$  oder  $\Im(u)$  ungleich dem Nullvektor. In ersterem Fall setze  $v_{n+1} = \Re(u) / \|\Re(u)\|$ , und analog im anderen.  $\square$

Nun soll Satz 7.33 übertragen werden auf beliebige normale Abbildungen.

**7.37 Satz** (Spektralsatz für normale Abbildungen) Sei  $V$  ein endlich dimensionaler komplexer Vektorraum und  $T \in \mathcal{L}(V)$  normal. Dann existiert eine Orthonormalbasis  $\mathcal{B} = (u_1, \dots, u_N)$ , so dass die darstellende Matrix  $T_{\mathcal{B}}^{\mathcal{B}}$  diagonal ist.

**Beweis:** Sei  $\mathcal{C}$  eine beliebige Orthonormalbasis von  $V$ . Dann ist  $T_{\mathcal{C}}^{\mathcal{C}} \in \text{Mat}(N \times N, \mathbb{C})$  normal, weil nach jeweils zweifacher Anwendung von Satz 7.29 und Satz 4.35(iv) gilt

$$(T_{\mathcal{C}}^{\mathcal{C}})^* T_{\mathcal{C}}^{\mathcal{C}} = (T^*)_{\mathcal{C}}^{\mathcal{C}} T_{\mathcal{C}}^{\mathcal{C}} = (T^* T)_{\mathcal{C}}^{\mathcal{C}} = (T T^*)_{\mathcal{C}}^{\mathcal{C}} = T_{\mathcal{C}}^{\mathcal{C}} (T^*)_{\mathcal{C}}^{\mathcal{C}} = T_{\mathcal{C}}^{\mathcal{C}} (T_{\mathcal{C}}^{\mathcal{C}})^* .$$

Nach Satz 7.33 existiert eine unitäre Matrix  $U \in \text{Mat}(N \times N, \mathbb{C})$ , so dass  $U^* T_{\mathcal{C}}^{\mathcal{C}} U = D$  diagonal ist. Somit gilt die Behauptung für  $\mathcal{B} = \mathcal{C}U$ .  $\square$

**7.38 Korollar** Für eine normale Abbildung auf einen endlich dimensional komplexen Vektorraum sind für alle Eigenwerte die geometrischen Vielfachheiten gleich den algebraischen Vielfachheiten.

Nun soll der Spektralsatz noch auf andere Art und Weise formuliert werden.

**7.39 Definition** Sei  $W \subset V$  ein Unterraum eines Vektorraumes  $V$  mit Skalarprodukt. Dann ist sein orthogonales Komplement definiert durch

$$W^{\perp} = \{v \in V \mid v \perp w \ \forall \ w \in W\} = \{v \in V \mid \langle v|w \rangle = 0 \ \forall \ w \in W\} .$$

Sei  $W' \subset V$  ein weiterer Unterraum. Dann sind  $W \perp W'$  orthogonal genau dann, wenn  $w \perp w'$  für alle  $w \in W$  und  $w' \in W'$ .

**7.40 Satz** Sei  $W$  ein Unterraum eines endlich dimensional Vektorraumes mit Skalarprodukt  $V$ . Dann gilt

- (i)  $W^{\perp}$  ist ein Unterraum.
- (ii)  $W \oplus W^{\perp} = V$ , d.h.  $W$  und  $W^{\perp}$  spannen  $V$  auf, aber  $W \cap W^{\perp} = \{0\}$ .
- (iii)  $(W^{\perp})^{\perp} = W$

Sei  $v \in V$  wie in (ii) zerlegt in  $v = w + u$  mit  $w \in W$  und  $u \in W^{\perp}$ . Diese Zerlegung ist eindeutig und definiert eine (dem Unterraum  $W$  zugehörige) Abbildung  $P : V \rightarrow V$  durch  $Pv = w$ , für welche gilt:

- (iv)  $P$  ist linear.
- (v)  $P^2 = P$ , d.h.  $P$  ist eine idempotente Abbildung (manchmal auch Projektion genannt).
- (vi)  $P^* = P$ , d.h.  $P$  ist eine Projektion (manchmal auch orthogonale Projektion genannt).
- (vii) Wenn  $u_1, \dots, u_M$  eine Orthonormalbasis von  $W$  ist, dann gilt

$$P = \sum_{m=1}^M |u_m\rangle \langle u_m| .$$

**Beweis:** (i) Seien  $v, v' \in W^{\perp}$  und  $\lambda \in \mathbb{K}$ . Dann gilt

$$\langle v + \lambda v' | w \rangle = \langle v | w \rangle + \lambda \langle v' | w \rangle = 0 + \lambda \cdot 0 = 0 ,$$

d.h.  $v + \lambda v' \in W^{\perp}$ . (ii) Sei  $u_1, \dots, u_M$  eine Orthonormalbasis von  $W$ , die nach dem Basiserweiterungssatz (Korollar 7.20) zu einer Basis  $u_1, \dots, u_N$  von  $V$  vervollständigt werden kann. Dann sind  $u_{M+1}, \dots, u_N$  in  $W^{\perp}$  und spannen  $W^{\perp}$  auch auf (Übung). (iii) folgt aus (ii). Für die Linearität von  $P$  in (iv), seien  $v, v' \in V$  und  $\lambda \in \mathbb{K}$ . Dann gilt  $v = w + u$  und  $v' = w' + u'$  mit  $w, w' \in W$  und  $u, u' \in W^{\perp}$ . Also sind  $w + \lambda w' \in W$  und  $u + \lambda u' \in W^{\perp}$  (weil  $W$  und  $W^{\perp}$  beides ja Unterräume sind) und  $v + \lambda v' = (w + \lambda w') + (u + \lambda u')$ .

Also ist  $P(v + \lambda v') = w + \lambda w' = Pv + \lambda Pv'$ . (v) ist offensichtlich nach der Definition der Abbildung  $P$ . Für (vi) gilt es zu zeigen, dass

$$\langle v | Pv' \rangle = \langle Pv | v' \rangle, \quad \forall v, v' \in V.$$

Hierfür werden  $v$  und  $v'$  nach der schon oben verwandten Orthonormalbasis entwickelt:

$$v = \sum_{n=1}^N v_n u_n, \quad v' = \sum_{n=1}^N v'_n u_n.$$

Dann gilt

$$Pv = \sum_{n=1}^M v_n u_n, \quad Pv' = \sum_{n=1}^M v'_n u_n,$$

und deswegen folgt die gewünschte Identität direkt durch Einsetzen. Der letzte Punkt (vii) wird genauso überprüft.  $\square$

**7.41 Bemerkungen** 1. Es gibt schon  $2 \times 2$ -Matrizen, die idempotent, aber nicht selbstadjungiert sind, z.B.

$$P = \begin{pmatrix} \frac{1}{2} - \sqrt{\frac{1}{4} - bc} & b \\ c & \frac{1}{2} + \sqrt{\frac{1}{4} - bc} \end{pmatrix}, \quad |bc| \leq \frac{1}{4}.$$

2. Gemäß der Definition von  $P$  gilt für  $w \in \text{Ran}(P)$  auch  $Pw = w$ . Tatsächlich ist dies richtig für jede idempotente Abbildung  $P^2 = P$ . In der Tat, sei  $w = Pv$ . Dann gilt

$$Pw - w = P^2v - Pv = Pv - Pv = 0,$$

was die Behauptung zeigt.  $\diamond$

**7.42 Satz** Seien  $W_1, \dots, W_K$  Unterräume eines endlich dimensionalen Vektorraumes  $V$  mit Skalarprodukt und seien  $P_1, \dots, P_K$  die zugehörigen Projektionen. Dann sind  $W_1, \dots, W_K$  paarweise orthogonal genau dann, wenn

$$P_k P_n = \delta_{k,n} P_k.$$

**Beweis:** "⇒" Wenn  $k = n$ , dann lautet die Gleichung  $(P_k)^2 = P_k$ , was gemäß Satz 7.40(v) für jede Projektion  $P_k$  gilt. Sei nun  $k \neq n$ . Sei  $u_1(k), \dots, u_{M(k)}(k)$  eine Orthonormalbasis von  $W_k$ . Unter Verwendung von Satz 7.40(vii) folgt dann

$$P_k P_n = \sum_{l=1}^{M(k)} |u_l(k)\rangle\langle u_l(k)| \sum_{l'=1}^{M(n)} |u_{l'}(n)\rangle\langle u_{l'}(n)| = \sum_{l=1}^{M(k)} \sum_{l'=1}^{M(n)} |u_l(k)\rangle\langle u_l(k)|u_{l'}(n)\rangle\langle u_{l'}(n)| = 0,$$

wobei Letzteres aus der Orthogonalität  $W_k \perp W_n$  folgt, die  $\langle u_l(k) | u_{l'}(n) \rangle = 0$  impliziert. "⇐" Seien  $k \neq n$  und  $w_k \in W_k$  und  $w_n \in W_n$ . Dann gilt  $P_k w_k = w_k$  und  $P_n w_n = w_n$ . Somit

$$\langle w_k | w_n \rangle = \langle P_k w_k | P_n w_n \rangle = \langle w_k | (P_k)^* P_n w_n \rangle = \langle w_k | P_k P_n w_n \rangle = 0,$$

wobei zunächst  $(P_k)^* = P_k$  verwandt wurde, und dann die Voraussetzung. Also  $W_k \perp W_n$ .  $\square$

Das orthogonale Komplement tritt auch in Zusammenhang mit Abbildungseigenschaften auf. Folgender Satz findet oft Anwendungen.



**7.43 Satz** Sei  $V$  ein Vektorraum mit Skalarprodukt und  $T \in \mathcal{L}(V)$ . Dann gilt

$$\text{Ker}(T^*) = \text{Ran}(T)^\perp.$$

**Beweis.** In der Tat:

$$\begin{aligned} v \in \text{Ker}(T^*) &\iff \langle w | T^* v \rangle = 0 \quad \forall w \in V \\ &\iff \langle T w | v \rangle = 0 \quad \forall w \in V \\ &\iff v \in \text{Ran}(T)^\perp, \end{aligned}$$

was den Beweis schon beendet. □

**7.44 Satz** (Projektionsoperatorversion des Spektralsatzes für normale Abbildungen) Sei  $V$  ein endlich dimensionaler komplexer Vektorraum mit Skalarprodukt und sei  $T \in \mathcal{L}(V)$ . Dann ist  $T$  normal genau dann, wenn Folgendes gilt: Wenn  $z_1, \dots, z_K \in \mathbb{C}$  die paarweise verschiedenen Eigenwerte von  $T$  sind und  $P_1, \dots, P_K$  die Projektionen auf die zugehörigen Eigenräume, dann gilt

$$T = \sum_{k=1}^K z_k P_k, \quad \mathbf{1}_V = \sum_{k=1}^K P_k, \quad P_k P_n = \delta_{k,n} P_k. \quad (7.4)$$

**Beweis:** Sei  $T$  normal und  $\mathcal{B} = (u_1, \dots, u_N)$  die in Satz 7.37 gegebene Orthonormalbasis, d.h.

$$T = (I_{\mathcal{B}})^{-1} D I_{\mathcal{B}}, \quad D = \begin{pmatrix} z'_1 & & \\ & \ddots & \\ & & z'_N \end{pmatrix} = \sum_{n=1}^N z'_n |e_n\rangle \langle e_n|,$$

wobei  $e_1, \dots, e_N$  die Standardbasis von  $\mathbb{C}^N$  ist. Somit gilt

$$T = \sum_{n=1}^N z'_n (I_{\mathcal{B}})^{-1} |e_n\rangle \langle e_n| I_{\mathcal{B}} = \sum_{k=1}^K z_k \sum_{n, z'_n=z_k} (I_{\mathcal{B}})^{-1} |e_n\rangle \langle e_n| I_{\mathcal{B}}.$$

Hier wurde die Summe umgeordnet, so dass lediglich paarweise verschiedene Eigenwerte  $z_k$  auftreten. Die Projektion auf den Eigenraum zu  $z_k$  ist aber gerade

$$P_k = \sum_{n, z'_n=z_k} (I_{\mathcal{B}})^{-1} |e_n\rangle \langle e_n| I_{\mathcal{B}},$$

und hieraus folgt die Darstellung von  $T$ , sowie die anderen Eigenschaften. Die Umkehrung, d.h. die Normalität von  $T$  gegeben durch (7.4), sei als Übung verifiziert. □

**7.45 Bemerkung** Mit Hilfe der Darstellung aus Satz 7.44 ist es sehr einfach sogenanntes Spektralkalkül zu machen, d.h. gegeben eine Funktion  $f : \mathbb{C} \rightarrow \mathbb{C}$ , kann eine Funktion  $f(T)$  von  $T$  definiert werden durch

$$f(T) = \sum_{k=1}^K f(z_k) P_k.$$

Dies ist wieder ein normaler Operator auf dem gleichen Vektorraum. Für eine analytische Funktion  $f$  (gegeben durch eine konvergente Potenzreihe) stimmt dies mit Definition 6.21 überein. Tatsächlich kann jede Funktion auf den endlich vielen Werten  $z_k$  durch ein Polynom dargestellt werden, so dass der Spektralsatz für normale lineare Abbildungen auf endlich dimensionalen Vektorräumen nicht erlaubt, weitere Funktionen der linearen Abbildung zu definieren. Auf unendlich dimensionalen Räumen hingegen können mit dem Spektralsatz für normale Abbildungen weitere Funktionen verwandt werden (z.B. Spektralkalkül mit stetigen Funktionen  $f$ ). ◇

## 8 Jordan'sche Normalform

Es sei daran erinnert, dass eine lineare Abbildung  $T$  auf einem Vektorraum  $V$  diagonalisierbar heißt genau dann, wenn es eine Basis  $\mathcal{B}$  von Eigenvektoren von  $V$  gibt. Dann ist die zugehörige darstellende Matrix  $T_{\mathcal{B}}^{\mathcal{B}}$  diagonal. Gemäß Satz 6.18 ist dies gleichbedeutend damit, dass die Summe  $\sum_{z_k} \beta(z_k)$  aller geometrischen Vielfachheiten über alle Eigenwerte  $z_k$  gleich der Dimension von  $V$  ist. Der Spektralsatz für normale Abbildungen besagt unter anderem, dass jede normale Abbildung diagonalisierbar ist, sogar mit einer Orthonormalbasis  $\mathcal{B}$ .

Nun gibt es lineare Abbildungen, die nicht diagonalisierbar sind. Dies tritt auch in komplexen Vektorräumen auf, in denen ja das charakteristische Polynom wegen des Fundamentalsatzes in Linearfaktoren zerfällt, und liegt dann daran, dass die geometrische Vielfachheit  $\beta(z)$  (mindestens) eines Eigenwertes  $z$  echt kleiner als die algebraische Vielfachheit  $\alpha(z)$  ist. Ein explizites Beispiel wurde schon in Bemerkung 6.25 angegeben, an das jetzt erinnert sei.

**8.1 Definition** Ein Jordanblock der Größe  $N$  mit Eigenwert  $z_1 \in \mathbb{C}$  ist folgende  $N \times N$  Matrix:

$$J_N(z_1) = (z_1 \delta_{n,m} + \delta_{n+1,m})_{n,m=1,\dots,N} = \begin{pmatrix} z_1 & 1 & 0 & \dots & 0 \\ 0 & z_1 & 1 & \dots & 0 \\ & & \ddots & \ddots & \vdots \\ & & & 0 & z_1 & 1 \\ 0 & & & & 0 & z_1 \end{pmatrix}.$$

Das charakteristische Polynom dieser Matrix ist  $p_{J_N(z_1)}(z) = (z_1 - z)^N$  und der Standardbasisvektor  $e_1$  ist der einzige Eigenvektor (bis auf Vielfache). Also gilt für die Vielfachheiten  $1 = \beta(z_1) < \alpha(z_1) = N$  für  $N \geq 2$ .

**8.2 Satz** (Jordan'sche Normalform) Sei  $T \in \mathcal{L}(V)$  eine lineare Abbildung auf einem endlich dimensionalen komplexen Vektorraum  $V$ . Dann gibt es eine sogenannte Jordan-Basis  $\mathcal{B}$  von  $V$  und Jordan Blöcke  $J_{N_1}(z'_1), \dots, J_{N_K}(z'_K)$ , so dass

$$T_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} J_{N_1}(z'_1) & & \\ & \ddots & \\ & & J_{N_K}(z'_K) \end{pmatrix}.$$

**8.3 Bemerkungen** 1. Wenn  $T$  durch eine Matrix  $A \in \text{Mat}(N \times N, \mathbb{C})$  gegeben ist, so besagt der Satz, dass es eine invertierbare komplexe Matrix  $M$  gibt, für welche

$$M^{-1}AM = \begin{pmatrix} J_{N_1}(z'_1) & & \\ & \ddots & \\ & & J_{N_K}(z'_K) \end{pmatrix},$$

und zwar  $M = (b_1, \dots, b_N)$ , wenn die im Satz auftretende geordnete Basis  $\mathcal{B} = (b_1, \dots, b_N)$  ist.

2. In der Jordan'schen Normalform können die Jordan Blöcke die gleichen Eigenwerte haben.
3. Wenn die Abbildung  $T$  auf einem reellen Vektorraum definiert ist (also  $\mathbb{R}$ -linear ist), aber zudem das charakteristische Polynom  $p_T$  über  $\mathbb{R}$  in Linearfaktoren zerfällt, so gilt der Satz auch und die  $z_k$

sind dann reell. Für reelle Matrizen  $A$  kann also unter dieser Zusatzannahme an das charakteristische Polynom eine reelle Matrix  $M$  mit obigen Eigenschaften gefunden werden. Für beliebige reelle Matrizen  $A$  (für die das charakteristische Polynom ggfs. nicht faktorisiert) kann immer die komplexe Jordan'schen Normalform bestimmt werden und aus dieser kann dann relativ leicht eine reelle Version abgeleitet werden (siehe Literatur).  $\diamond$

Der Beweis des Satzes wird den Großteil dieses Kapitels einnehmen und er wird auch aufzeigen, wie die Jordanbasis konstruiert werden kann. Zunächst werden folgende Konzepte eingeführt.

**8.4 Definition** Sei  $V$  ein Vektorraum und  $T \in \mathcal{L}(V)$  mit Eigenwert  $z$ . Der verallgemeinerte Eigenraum von  $T$  zu  $z$  der Stufe  $j \geq 1$  ist

$$E_{T,j}(z) = \text{Ker}((T - z\mathbf{1})^j).$$

Vektoren aus  $E_{T,j}(z)$  heißen auch verallgemeinerte Eigenvektoren von  $T$  zum Eigenwert  $z$ . Der Hauptraum von  $T$  zu  $z$  ist  $H_T(z) = \cup_{j \geq 1} E_{T,j}(z)$ .

**8.5 Definition** Sei  $V$  ein Vektorraum und  $S \in \mathcal{L}(V)$ . Dann heißt  $S$  nilpotent genau dann, wenn es ein  $n \in \mathbb{N}$  gibt mit  $S^n = 0$ .

**8.6 Bemerkungen** 1. Als Kern einer linearen Abbildung sind die  $E_{T,j}(z)$  alle Unterräume. Der Eigenraum  $E_T(z)$  von  $T$  zu  $z$  ist gleich  $E_{T,1}(z)$ . Offensichtlich gilt  $E_{T,j}(z) \subset E_{T,j+1}(z)$ . Andererseits können diese Vektorräume bei einem endlich dimensionalen  $V$  auch nicht beliebig groß werden.

2. Für einen Jordanblock  $T = J_N(z)$  gilt

$$E_{T,j}(z) = \text{span}(e_1, \dots, e_j), \quad j = 1, \dots, N.$$

In der Tat,

$$(T - z\mathbf{1})^j = \begin{pmatrix} 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & 0 & \\ & \ddots & \ddots & \ddots & \\ & & 0 & 0 & 1 \\ 0 & & & 0 & 0 \end{pmatrix}^j = \begin{pmatrix} \dots & 0 & 1 & 0 & \dots \\ & \ddots & \ddots & \ddots & \\ & & 0 & 0 & 1 \\ & & & 0 & 0 \\ 0 & & & 0 & \vdots \end{pmatrix},$$

wobei die ersten  $j$  Spalten gleich Nullvektoren sind. Der Hauptraum ist  $H_T(z) = \mathbb{K}^N$ . Außerdem ist

$$J_z(0) = J_N(z) - z\mathbf{1}_N = T - z\mathbf{1}_N$$

nilpotent.  $\diamond$

**8.7 Lemma** (Lemma von Fitting) Sei  $V$  ein endlich dimensionaler Vektorraum über  $\mathbb{K}$ . Weiter sei  $T \in \mathcal{L}(V)$  mit Eigenwert  $z \in \mathbb{K}$  und verallgemeinerten Eigenräumen  $E_j = E_{T,j}(z) = \text{Ker}((T - z\mathbf{1})^j)$ . Definiere den Fitting Index von  $T$  zu  $z$  als

$$f = \min\{j \in \mathbb{N} \mid E_{j+1} = E_j\},$$

und setze  $R_j = \text{Ran}((T - z\mathbf{1})^j)$ . Dann gilt

$$(i) \quad f = \min\{j \in \mathbb{N} \mid R_j = R_{j+1}\}$$

(ii)  $T : R_f \rightarrow R_f$  ist ein Isomorphismus.

(iii)  $E_{f+i} = E_f$  und  $R_{f+i} = R_f$  für alle  $i \geq 0$ . Insbesondere  $E_f = H_T(z)$ .

(iv)  $(T - z\mathbf{1})^f(E_f) = \{0\}$ , d.h.  $T - z\mathbf{1}$  ist nilpotent.

(v)  $E_f$  und  $R_f$  sind invariant unter  $T$ , d.h.  $T(E_f) \subset E_f$  und  $T(R_f) \subset R_f$ .

(vi)  $V = E_f \oplus R_f$

(vii)  $\dim(E_f)$  ist gleich der algebraischen Multiplizität von  $z$ .

**Beweis:** Es darf  $z = 0$  angenommen werden (ersetze  $T - z\mathbf{1}$  durch  $T$ ). Schon oben in der Bemerkung wurde auf  $E_j \subset E_{j+1}$  hingewiesen. Zudem gilt  $R_j \supset R_{j+1}$ , da  $R_{j+1} = T^{j+1}V = T^j R_1$ . Somit erhält man folgendes Diagramm:

$$\begin{array}{ccc} E_j \subset V & \xrightarrow{T^j} & R_j \\ \cap & \parallel & \cup \\ E_{j+1} \subset V & \xrightarrow{T^{j+1}} & R_{j+1} \end{array}$$

Nun impliziert der Dimensionssatz 4.8 angewandt auf  $T^j : V \rightarrow V$  und  $T^{j+1} : V \rightarrow V$ , dass

$$\dim(V) = \dim(E_j) + \dim(R_j) = \dim(E_{j+1}) + \dim(R_{j+1}).$$

Somit folgt

$$\begin{aligned} R_{j+1} = R_j &\iff \dim(R_{j+1}) = \dim(R_j) && \text{(rechte Spalte des Diagramms)} \\ &\iff \dim(E_{j+1}) = \dim(E_j) && \text{(Dimensionssatz)} \\ &\iff E_{j+1} = E_j && \text{(linke Spalte des Diagramms)} \end{aligned}$$

Dies impliziert (i). Außerdem folgt aus  $R_f = R_{f+1}$ , dass  $T : R_f \rightarrow R_{f+1} = R_f$  surjektiv und somit ein Isomorphismus ist, was (ii) ist. Daraus folgt dann  $R_{f+i} = R_f$  für alle  $i \geq 1$ , und unter Verwendung der Äquivalenz ist auch (iii) nun bewiesen. (iv) ist offensichtlich. In (v) fehlt lediglich  $TE_f \subset E_{f-1} \subset E_f$ . Nun zu (vi). Nach dem Dimensionssatz gilt  $\dim(E_f) + \dim(R_f) = \dim(V)$ . Es verbleibt also zu zeigen, dass  $E_f \cap R_f = \{0\}$ . Sei  $v \in E_f \cap R_f$ . Dann gilt  $T^f v = 0$  und  $v = T^f w$  für ein  $w \in V$ . Somit ist  $T^{2f} w = 0$ , d.h.  $w \in E_{2f} = E_f$ . Also  $v = T^f w = 0$ . Letztendlich kommen wir zu (vii). Sei  $d = \dim(E_f)$ . Die Abbildung  $S = T|_{E_f}$  erfüllt dann  $S^f = 0$  nach (iv). Also kann sie nur 0 als Eigenwert haben und das charakteristische Polynom ist  $p_S(z) = (-z)^d$ , da es ja vom Grade  $d$  ist und keinen weiteren Eigenwert haben kann. Sei nun  $\mathcal{B} = (\mathcal{B}', \mathcal{B}'')$  eine Basis von  $V$  bestehend aus einer Basis  $\mathcal{B}'$  von  $E_f$  und einer Basis  $\mathcal{B}''$  von  $R_f$ . Nach (v) gilt nun für die darstellende Matrix

$$T_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} (T|_{E_f})_{\mathcal{B}'}^{\mathcal{B}'} & 0 \\ 0 & (T|_{R_f})_{\mathcal{B}''}^{\mathcal{B}''} \end{pmatrix} = \begin{pmatrix} S_{\mathcal{B}'}^{\mathcal{B}'} & 0 \\ 0 & (T|_{R_f})_{\mathcal{B}''}^{\mathcal{B}''} \end{pmatrix}.$$

Also folgt für das charakteristische Polynom

$$p_T(z) = p_S(z) p_{T|_{R_f}}(z) = (-z)^d p_{T|_{R_f}}(z).$$

Nun ist aber  $T|_{R_f}$  ein Isomorphismus nach (ii) und somit ist  $p_{T|_{R_f}}(0) \neq 0$ . Also ist  $d$  gleich der algebraischen Multiplizität des Eigenwertes 0.  $\square$

**8.8 Satz** (Hauptraumzerlegung) Sei  $T \in \mathcal{L}(V)$  eine lineare Abbildung auf einem endlich dimensionalen Vektorraum  $V$  über  $\mathbb{K}$ . Es gelte folgende Faktorisierung des charakteristischen Polynomes

$$p_T(z) = (z_1 - z)^{\alpha(z_1)} \cdots (z_K - z)^{\alpha(z_K)} .$$

Dann gelten für die Haupträume  $H_k = H_T(z_k)$ ,  $k = 1, \dots, K$ , mit  $\dim(H_k) = \alpha(z_k)$  folgende Aussagen:

(i)  $T(H_k) \subset H_k$ , d.h. die  $H_k$  sind  $T$ -invariante Unterräume.

(ii)  $V = H_1 \oplus \dots \oplus H_K$

(iii) Es gibt eine Basis  $\mathcal{B}$  von  $V$ , so dass

$$T_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} z_1 \mathbf{1}_{\alpha(z_1)} + S_1 & & \\ & \ddots & \\ & & z_K \mathbf{1}_{\alpha(z_K)} + S_K \end{pmatrix} ,$$

wobei  $S_k : \mathbb{K}^{\alpha(z_k)} \rightarrow \mathbb{K}^{\alpha(z_k)}$  nilpotent ist.

(iv)  $T = D + S$ , wobei  $D$  diagonalisierbar ist,  $S$  nilpotent und  $DS = SD$ .

**Beweis:** Die beiden ersten Aussagen folgen direkt nach iterativer Anwendung des Lemmas von Fitting (Induktion über die Anzahl  $K$  von Eigenwerten). Die Basis in (iii) ist gegeben durch  $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_K)$ , wobei  $\mathcal{B}_k$  eine Basis von  $H_k$  ist. Hierbei ist die Zerlegung von  $T$  auf  $H_k$  durch Lemmas 8.7(iv) gegeben. Die Zerlegung in (iv) ist nach (iii) klar und  $DS = SD$  kann nachgerechnet werden.  $\square$

Um den Beweis von Satz 8.2 zu vervollständigen, müssen jetzt noch die in der Hauptraumzerlegung auftretenden nilpotenten Matrizen durch adäquate Wahl der Basis in eine kanonische Form gebracht werden, nämlich in eine Summe von Jordanblöcken. Dass dies möglich ist, zeigt der folgende Satz.

**8.9 Satz** (Jordan'sche Normalform für eine nilpotente Abbildung) Sei  $S \in \mathcal{L}(V)$  eine nilpotente lineare Abbildung auf einem endlich dimensionalen Vektorraum  $V$  über  $\mathbb{K}$ . Dann gibt es eine Basis  $\mathcal{B}$  von  $V$  und  $n_1, \dots, n_K \in \mathbb{N}$  mit  $n_1 + \dots + n_K = \dim(V)$ , so dass

$$S_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} J_{n_1}(0) & & \\ & \ddots & \\ & & J_{n_K}(0) \end{pmatrix} .$$

**Beweis:** Sei  $E_j = \text{Ker}(S^j)$ . Dann erhält man folgende aufsteigende Folge von Unterräumen:

$$\{0\} = E_0 \subset E_1 \subset E_2 \dots \subset E_{f-1} \subset E_f = V ,$$

wobei  $f$  der Fitting Index von  $S$  ist (zum einzigen Eigenwert 0). Hieraus folgt (gemäß der Definition des Fitting Indizes), dass alle Inklusionen echt sind. Außerdem gilt wegen der Definition der  $E_j$

$$S(E_j) \subset E_{j-1} , \quad j = 1, \dots, f . \tag{8.1}$$

Des Weiteren wird Folgendes mehrfach verwandt werden:

**Fakt:** Sei  $j \geq 1$ . Wenn  $v_1, \dots, v_R \in E_{j+1}$  linear unabhängig sind und  $\text{span}\{v_1, \dots, v_R\} \cap E_j = \{0\}$ , so sind auch  $Sv_1, \dots, Sv_R \in E_j$  linear unabhängig und  $\text{span}\{Sv_1, \dots, Sv_R\} \cap E_{j-1} = \{0\}$ .

**Begründung:** Zunächst sind in der Tat  $Sv_1, \dots, Sv_R \in E_j$  wegen (8.1). Seien nun  $\lambda_1, \dots, \lambda_R \in \mathbb{K}$ , so dass

$$0 = \sum_{r=1}^R \lambda_r Sv_r = S \left( \sum_{r=1}^R \lambda_r v_r \right).$$

Nun ist entweder  $w = \sum_{r=1}^R \lambda_r v_r$  gleich dem Nullvektor, oder aber wegen  $\text{span}\{v_1, \dots, v_R\} \cap E_j = \{0\}$  gilt  $w \notin E_j$  und wegen  $E_1 \subset E_j$  somit  $w \notin E_1 = \text{Ker}(S)$ , woraus wieder  $w = 0$  folgt. Nun impliziert die lineare Unabhängigkeit der  $v_1, \dots, v_R$ , deswegen  $\lambda_1 = \dots = \lambda_R = 0$ . Also sind  $Sv_1, \dots, Sv_R$  linear unabhängig. Es verbleibt noch die letzte Eigenschaft zu zeigen. Sei  $u \in \text{span}\{Sv_1, \dots, Sv_R\} \cap E_{j-1}$ , d.h.  $u = \sum_{r=1}^R \lambda_r Sv_r = Sw$  und  $S^{j-1}u = 0$  (mit  $w = \sum_{r=1}^R \lambda_r v_r$  wie oben). Also  $S^j w = 0$ , so dass  $w \in E_j$  und nach Annahme also  $w = 0$ . Daraus folgt auch  $u = Sw = 0$ .  $\diamond$

Eine erste Folgerung aus dem Fakt betreffen die natürlichen Zahlen

$$r_j = \dim(E_j) - \dim(E_{j-1}).$$

Sie erfüllen die Ungleichungen

$$r_f \leq r_{f-1} \leq r_{f-2} \leq \dots \leq r_1.$$

In der Tat,  $r_{j+1}$  ist genau gleich der maximalen Anzahl linear unabhängiger Vektoren in  $E_{j+1}$  derart, dass ihr Spann mit  $E_j$  einen trivialen Schnitt hat. Wird der Fakt mit  $R = r_{j+1}$  auf eine spezifische Wahl dieser Vektoren angewandt, so erhält man  $R$  linear unabhängige Vektoren in  $E_j \setminus E_{j-1}$ . Die maximale Anzahl  $r_j$  linear unabhängiger Vektoren in  $E_j \setminus E_{j-1}$  ist also größer oder gleich  $R$ , d.h.  $r_j \geq R = r_{j+1}$ .

Nun können wir die Konstruktion der Jordanbasis beginnen. Zunächst werden  $r_f$  linear unabhängige Vektoren  $b_1, \dots, b_{r_f}$  gewählt, so dass  $\text{span}\{b_1, \dots, b_{r_f}\} \cap E_{f-1} = \{0\}$  (dies ist immer möglich, denn man wählt den  $r_f$  dimensionalen Unterraum  $\text{span}\{b_1, \dots, b_{r_f}\}$ , so dass er den Unterraum  $E_{f-1}$  der Dimension im  $\dim(V) - r_f$  nur trivial schneidet). Für jedes  $r = 1, \dots, r_f$  erhält man hieraus eine sogenannte Jordankette

$$b_r, Sb_r, \dots, S^{f-1}b_r.$$

Jede dieser Ketten spannt einen  $f$ -dimensionalen  $S$ -invarianten Unterraum  $I_r^{(f)}$  auf, und auf diesem Unterraum ist die darstellende Matrix bez. der geordneten Basis  $\mathcal{B}_r = (S^{f-1}b_r, \dots, Sb_r, b_r)$  tatsächlich ein Jordanblock:

$$(S|_{I_r^{(f)}})_{\mathcal{B}_r} = J_f(0),$$

denn die Wirkung von  $S$  ist nur ein Shift (Weiterschieben) der Basisvektoren:

$$S\mathcal{B}_r = (0, S^{f-1}b_r, \dots, S^2b_r, Sb_r).$$

Außerdem sind diese  $r_f$  Jordanketten der Länge  $f$  alle auf disjunkten Unterräumen definiert, denn obiger Fakt iterativ angewandt impliziert ja, dass

$$b_1, \dots, b_{r_f}, Sb_1, \dots, Sb_{r_f}, \dots, S^{f-1}b_1, \dots, S^{f-1}b_{r_f}$$

alle linear unabhängig sind und somit einen Unterraum der Dimension  $fr_f$  aufspannen, der genau  $r_f$  Jordanblöcke der Größe  $f$  enthält. Wenn

$$I^{(f)} = I_1^{(f)} \oplus \dots \oplus I_{r_f}^{(f)},$$

und als Basis von  $I^{(f)}$  verwandt wird

$$\mathcal{B}^{(f)} = (S^{f-1}b_1, \dots, Sb_1, b_1, S^{f-1}b_2, \dots, Sb_2, b_2, \dots, S^{f-1}b_{r_f}, \dots, Sb_{r_f}, b_{r_f}), \quad (8.2)$$

so gilt

$$(S|_{I^{(f)}})_{\mathcal{B}^{(f)}}^{\mathcal{B}^{(f)}} = \begin{pmatrix} J_f(0) & & \\ & \ddots & \\ & & J_f(0) \end{pmatrix},$$

mit  $r_f$  Jordanblöcken auf der Diagonale.

Falls nun  $r_f = r_1$  gilt, so ist der Beweis beendet, weil  $\dim(V) = fr_f$  gleich der Anzahl der Elemente von  $\mathcal{B}^{(f)}$  ist. Falls  $r_f > r_1$ , so gibt es ein  $j$  mit  $r_f = r_{j+1} < r_j$ . Nun können die oben konstruierten Vektoren  $S^{r_f-j}b_1, \dots, S^{r_f-j}b_{r_f} \in E_j \setminus E_{j-1}$  noch durch  $s_j = r_{j-1} - r_j$  weitere Vektoren  $b_{r_{j+1}}, \dots, b_{r_j} \in E_j \setminus E_{j-1}$  ergänzt werden, so dass all diese Vektoren in  $E_j \setminus E_{j-1}$  linear unabhängig sind und der von ihnen aufgespannte Unterraum nur einen trivialen Schnitt mit  $E_{j-1}$  hat. Nun erhält man  $s_j$  zusätzliche Jordanketten der Länge  $j$ :

$$b_{r_{j+1}+s}, Sb_{r_{j+1}+s}, \dots, S^{j-1}b_{r_{j+1}+s}, \quad s = 1, \dots, s_j.$$

Wiederum garantiert obiger Fakt, dass all diese Vektoren linear unabhängig voneinander, aber auch von all den vorher konstruierten sind. Wenn diese Vektoren zur obigen Basis von  $I^{(f)}$  hinzugefügt werden, entstehen  $s_j$  Jordanblöcke  $J_j(0)$ . Nun wird diese Prozedur endlich oft wiederholt. Am Ende bleibt evtl.  $s_1 = r_1 - r_2 > 0$ . Die als Letztes hinzugefügten Vektoren bilden dann Jordanketten der Länge 1, sind also Eigenvektoren, die nicht Bild von verallgemeinerten Eigenvektoren sind.  $\square$

Es gibt Beweise von Satz 8.9, die kürzer erscheinen mögen, die aber nicht immer so konstruktiv sind. Die Beweisschritte im obigen Beweis können in der expliziten Konstruktion einer Jordanbasis zu einer nilpotenten Abbildung direkt verwandt werden (Kochrezept):

- Berechne die Potenzen  $S^j$ .
- Berechne  $E_j = \ker(S^j)$  mit Hilfe des Gauss Algorithmus.
- Wähle sukzessive zusätzliche linear unabhängige Vektoren in  $E_j \setminus E_{j-1}$ , die noch nicht im Spann längerer Jordan-Ketten liegen und deren Spann  $E_{j-1}$  nur trivial schneidet.
- Bilde die zugehörigen Jordanketten und bilde die Jordanbasis in der Reihenfolge wie in (8.2), d.h. es wird in untenstehender Graphik umgekehrte Zeile nach umgekehrter Zeile aneinandergesetzt.

Graphisch kann diese Prozedur wie folgt dargestellt werden. Unten wird dies an einem Beispiel illustriert.

$$\begin{array}{ccccccc}
 E_f \setminus E_{f-1} & \xrightarrow{S} & E_{f-1} \setminus E_{f-2} & \xrightarrow{S} & E_{f-2} \setminus E_{f-3} & \xrightarrow{S} & \dots \xrightarrow{S} E_2 \setminus E_1 \xrightarrow{S} E_1 \\
 \hline
 b_1 & & Sb_1 & & S^2b_1 & & S^{f-2}b_1 & & S^{f-1}b_1 \\
 b_2 & & Sb_2 & & S^2b_2 & & S^{f-2}b_2 & & S^{f-1}b_2 \\
 \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
 b_{r_f} & & Sb_{r_f} & & S^2b_{r_f} & & S^{f-2}b_{r_f} & & S^{f-1}b_{r_f} \\
 & & b_{r_f+1} & & Sb_{r_f+1} & & S^{f-3}b_{r_f+1} & & S^{f-2}b_{r_f+1} \\
 & & \vdots & & \vdots & & \vdots & & \vdots \\
 & & b_{r_f-1} & & Sb_{r_f-1} & & S^{f-3}b_{r_f-1} & & S^{f-2}b_{r_f-1} \\
 & & & & \dots & & & & \\
 & & & & & & \dots & & \\
 & & & & & & b_{r_3+1} & & Sb_{r_3+1} \\
 & & & & & & \vdots & & \vdots \\
 & & & & & & b_{r_2} & & Sb_{r_2} \\
 & & & & & & & & b_{r_2+1} \\
 & & & & & & & & \vdots \\
 & & & & & & & & b_{r_1}
 \end{array}$$

- 8.10 Bemerkungen**
1. Außerdem liefert der Beweis genaue Informationen darüber, wie viele Jordanblöcke einer vorgegebenen Größe vorliegen. Z.B. ist die Größe der größten Jordanblöcke durch den Fitting Index  $f$  gegeben und die Anzahl dieser Blöcke durch  $\dim(V) - \dim(\text{Ker}(S^{f-1}))$  (was oben mit  $r_f$  bezeichnet ist).
  2. Der Beweis zeigt auch, dass sehr viele Freiheiten bei der Wahl der Jordanbasis bestehen. Selbst wenn nur ein einziger Jordanblock vorliegt, darf der Vektor  $b \in E_f \setminus E_{f-1}$  lediglich nicht in der Hyperebene  $E_{f-1}$  von Kodimension 1 liegen. Wenn allerdings  $b$  gewählt ist, so liegen  $Sb, \dots, S^{f-1}b$  schon fest.  $\diamond$

**8.11 Beispiel** Sei

$$S = \begin{pmatrix} 0 & 1 & 0 & 1 \\ -1 & 1 & 1 & 0 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & -1 & 0 \end{pmatrix},$$

Dann ist

$$S^2 = \begin{pmatrix} 0 & 1 & 0 & 1 \\ -1 & 1 & 1 & 0 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 1 \\ -1 & 1 & 1 & 0 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Also ist  $S$  nilpotent und sein Fitting Index ist 2. Es muss lediglich der Kern von  $S$  berechnet werden, z.B. mit dem Gauss-Algorithmus:

$$E_1 = \text{Ker}(S) = \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \end{pmatrix} \right\}.$$



Da  $\dim(E_2) = \dim(\mathbb{C}^4) = 4$  und  $\dim(E_1) = 2$  ist, gibt es  $r - f = 4 - 2 = 2$  Jordanblöcke der Größe  $f = 2$ . Da sich deren Dimensionen zu  $2 + 2 = 4$  aufaddieren, kann es keine weiteren Jordanblöcke geben. Eine mögliche Wahl der zwei Vektoren in  $E_2 \setminus E_1 = \mathbb{C}^4 \setminus E_1$  ist

$$b_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

Dann sind

$$Sb_1 = \begin{pmatrix} 0 \\ -1 \\ 1 \\ 1 \end{pmatrix}, \quad Sb_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \end{pmatrix}.$$

Somit ist ein gesuchter Basiswechsel

$$M = \begin{pmatrix} 0 & 1 & 1 & 0 \\ -1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 \end{pmatrix}.$$

Dann folgt ohne weitere Rechnung aus der Konstruktion

$$M^{-1}SM = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Natürlich kann dies auch noch einmal explizit durch Berechnung der Matrixmultiplikation überprüft werden.  $\diamond$

### 8.12 Beispiel Sei

$$S = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & a \\ 0 & 0 & 0 \end{pmatrix}, \quad a \in \mathbb{C}.$$

Es soll eine Jordan-Basis bestimmt werden in Abhängigkeit von  $a$ . Zunächst ist  $S$  nilpotent:

$$S^2 = \begin{pmatrix} 0 & 0 & a \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad S^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Es ist eine Fallunterscheidung notwendig. Zunächst sei  $a \neq 0$ . Dann ist der Fitting Index von  $S$  gleich 3. Also gibt es eine Jordankette der Länge 3 und aus Dimensionsgründen keine weitere. Nun ist

$$E_2 = \text{Ker}(S^2) = \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Also kann gewählt werden

$$b_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \in E_3 \setminus E_2.$$

Dann sind

$$Sb_1 = \begin{pmatrix} 1 \\ a \\ 0 \end{pmatrix}, \quad S^2b_1 = \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix}.$$

Somit führt

$$M = \begin{pmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

also zu

$$M^{-1}SM = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Nun wird der Fall  $a = 0$  betrachtet. Dann ist der Fitting Index  $f = 2$ . Es gibt also eine Jordan Kette der Länge 2 (und tatsächlich aus Dimensionsgründen nicht mehr als eine), und des weiteren eine Kette der Länge 1. Nun ist  $E_2 = \mathbb{C}^3$  und

$$E_1 = \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\}.$$

Also ist es möglich,

$$b'_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \in E_2 \setminus E_1$$

zu wählen. Dann ist

$$Sb'_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Zuletzt wählen wir ein  $b'_2 \in E_1$  linear unabhängig von  $Sb'_1$  als

$$b'_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Also gilt für  $a = 0$

$$(M')^{-1}SM' = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad M' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix},$$

was wiederum direkt verifiziert werden kann. ◇

Nun können wir die Sätze zusammenfügen, um den Hauptsatz zu beweisen.

**Beweis** von Satz 8.2: Wenn  $\mathbb{K} = \mathbb{C}$  ist, so impliziert der Fundamentalsatz die Faktorisierung des charakteristischen Polynoms, die in Satz 8.8 zur Hauptraumerlegung vorausgesetzt wird. Wenn nun in jedem Hauptraum  $H_k$  auf die nilpotente Abbildung  $S_k$  darauf Satz 8.9 angewandt wird, so erhält man eine Jordanbasis. □

Zusammenfassend haben wir also folgende Prozedur zur Konstruktion einer Jordanbasis zu einer linearen Abbildung  $T$  (Kochrezept):

- Berechne das charakteristische Polynom  $p_T$  und faktorisier es über  $\mathbb{C}$  (evtl. numerisch).
- Zu jedem Eigenwert  $z_j$  bestimme den Hauptraum  $H_T(z_j) = \text{Ker}((T - z_j \mathbf{1})^k)$ , wobei  $k$  ausreichend groß sein muss (mindestens gleich dem Fitting Index zu  $z_j$ ).
- Innerhalb jedes Hauptraumes bestimme eine Jordanbasis zur nilpotenten Abbildung  $(T - z_j \mathbf{1})|_{H_T(z_j)}$  gemäß oben beschriebener Vorgehensweise.

Dies sei wieder an einem Beispiel illustriert.

### 8.13 Beispiel Sei

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & -1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Es soll wieder eine Jordan-Basis bestimmt werden. Zunächst ist das charakteristische Polynom  $p_A(z) = (2 - z)(1 - z)^2$ . Also sind  $z_1 = 2$  und  $z_2 = 1$  die Eigenwerte. Da  $z_1 = 2$  algebraische Vielfachheit 1 hat, ist der Hauptraum gleich dem Eigenraum und aufgespannt von einem Eigenvektor  $b_1$  gegeben durch

$$E_{A,1}(2) = \text{Ker}(A - 2 \mathbf{1}) = \text{Ker} \begin{pmatrix} -1 & 0 & 0 \\ 1 & -1 & -1 \\ 0 & 0 & 0 \end{pmatrix} = \text{span} \left\{ \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\}.$$

Hingegen hat  $z_2$  algebraische Vielfachheit 2, und somit kann es einen nicht-diagonalen Jordanblock geben. Zunächst werden die Eigenvektoren berechnet:

$$E_{A,1}(1) = \text{Ker}(A - 1 \mathbf{1}) = \text{Ker} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix} = \text{span} \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Da dies nur 1-dimensional ist (geometrische Vielfachheit gleich 1), muss der Hauptraum  $H_A(1)$  größer sein. Hier kann er aus Dimensionsgründen höchstens von Dimension 2 sein. Also ist der Fitting Index zu  $z_2 = 1$  gleich 2. Es gilt

$$H_A(1) = E_{A,2}(1) = \text{Ker}((A - 1 \mathbf{1})^2) = \text{Ker} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \text{span} \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

Tatsächlich kann überprüft werden  $E_{A,3}(1) = E_{A,2}(1)$ . Nun wählen wir

$$b_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \in E_{A,2}(1) \setminus E_{A,1}(1).$$

Dann

$$(A - \mathbf{1})b_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Somit führt

$$M = (b_1, (A - \mathbf{1})b_2, b_2) = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

also zu

$$M^{-1}AM = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

was wiederum direkt überprüft werden sollte. ◇

Es sei nun eine der Standardanwendungen der Jordan'schen Normalform beschrieben werden, nämlich das Lösen linearer Differentialgleichungssysteme erster Ordnung mit konstanten Koeffizienten. Letztere sind von der Form

$$\dot{x}(t) = Ax(t), \quad x(0) = x_0,$$

wobei  $A \in \text{Mat}(N \times N, \mathbb{C})$  und  $t \in \mathbb{R} \mapsto x(t) \in \mathbb{C}^N$  differenzierbar ist mit Ableitung  $\dot{x}(t) = \partial_t x(t)$  und  $x_0 \in \mathbb{C}^N$  eine sogenannte Anfangsbedingung. Eine Lösung ist gegeben durch

$$x(t) = e^{tA} x_0,$$

wobei  $e^{At} = \sum_{k \geq 0} \frac{1}{k!} t^k A^k$  die Exponentialfunktion der Matrix  $tA$  ist. In der Tat, vertauschen von Summe und Ableitung (erlaubt wegen gleichmäßiger Konvergenz)

$$\dot{x}(t) = \partial_t \sum_{k \geq 0} \frac{t^k}{k!} A^k x_0 = \sum_{k \geq 1} \frac{\partial_t t^k}{k!} A^k x_0 = \sum_{k \geq 1} \frac{k t^{k-1}}{k!} A^k x_0 = A \sum_{k \geq 0} \frac{t^k}{k!} A^k x_0 = Ax(t).$$

Außerdem ist es möglich zu beweisen, dass die Lösung eindeutig ist, entweder mit dem Satz von Picard-Lindelöf (Vorlesung über gewöhnliche Differentialgleichungen) oder direkt (bilde  $y(t) = e^{-tA}x(t)$  und zeige, dass  $\dot{y}(t) = 0$  so dass  $y(t) = \text{konst}$ , und diese Konstante kann bei  $t = 0$  berechnet werden). Somit ist die Berechnung der Lösung der Differentialgleichung auf die Berechnung der Exponentialfunktion  $e^{tA}$  reduziert.

Falls  $A$  diagonalisierbar ist, so wurde dies schon in Satz 6.22 durchgeführt. Wenn  $A$  nicht diagonalisierbar ist, so gibt es zumindestens die (evtl. komplexe) Jordan'sche Normalform, d.h. es existiert eine invertierbare Matrix  $M$ , so dass

$$A = M^{-1}(D + J)M, \quad [D, J] = DJ - JD = 0,$$

wobei  $D$  diagonal ist und  $J$  eine nilpotente Matrix in Jordan'scher Normalform ist (genau wie in Satz 8.9 gegeben). Nun gilt

$$e^{tA} = \sum_{k \geq 0} \frac{t^k}{k!} A^k = \sum_{k \geq 0} \frac{t^k}{k!} (M^{-1}(D + J)M)^k = M^{-1} \sum_{k \geq 0} \frac{t^k}{k!} (D + J)^k M = M^{-1} e^{t(D+J)} M.$$

Tatsächlich gilt nach gleicher Rechnung für jede analytische Funktion  $f$  (konvergente Potenzreihe), so dass  $f(A) = M^{-1}f(D + J)M$ . Außerdem kommutieren  $D$  und  $J$ . Somit kann die Rechnung, die zu  $e^{a+b} = e^a e^b$  für  $a, b \in \mathbb{C}$  führt, direkt übertragen werden, um zu zeigen, dass

$$e^{tA} = M^{-1} e^{tD} e^{tJ} M. \tag{8.3}$$

Hier ist nun nicht nur  $e^{tD}$  leicht berechenbar, sondern auch die block-diagonale Matrix  $J$ , weil nämlich für jeden ihrer Diagonaleinträge gilt:

$$e^{tJ_n(0)} = \begin{pmatrix} 1 & \frac{t}{1!} & \frac{t^2}{2!} & \dots & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & \frac{t}{1!} & \dots & \\ & \ddots & \ddots & \ddots & \\ & & 0 & 1 & \frac{t}{1!} \\ 0 & & & 0 & 1 \end{pmatrix}, \quad J_n(0) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ & \ddots & \ddots & \ddots & \\ & & 0 & 0 & 1 \\ 0 & & & 0 & 0 \end{pmatrix}.$$

Dies wurde schon in den Übungen nachgerechnet. Somit sind nun alle Faktoren in (8.3) im Prinzip berechenbar. Dies sei nun an einem Beispiel illustriert.

**8.14 Beispiel** Es soll die Lösung des Anfangswertproblems

$$\dot{x}(t) = Ax(t), \quad A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & -1 \\ 0 & 0 & 2 \end{pmatrix}, \quad x(0) = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix},$$

berechnet werden. In Beispiel 8.13 wurde schon gezeigt, dass

$$\begin{aligned} A &= \begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \left[ \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \right] \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Also

$$e^{tA} = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} e^{2t} & 0 & 0 \\ 0 & e^t & 0 \\ 0 & 0 & e^t \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} e^t & 0 & 0 \\ te^t & e^t & e^t - e^{2t} \\ 0 & 0 & e^{2t} \end{pmatrix},$$

und die Lösung ist

$$x(t) = e^{tA} x(0) = \begin{pmatrix} e^t \\ te^t - e^{2t} \\ e^{2t} \end{pmatrix},$$

was durch Einsetzen nochmals überprüft werden kann. ◇

Nun wird noch eine zweite Anwendung der Jordan'schen Normalform präsentiert. Zunächst sei an den Begriff des Spektrums einer Matrix  $A \in \text{Mat}(N \times N, \mathbb{C})$  erinnert (Definition 6.31 und Satz 6.32):

$$\text{Spec}(A) = \{z \in \mathbb{C} \mid z \text{ Eigenwert von } A\} = \{z \in \mathbb{C} \mid z\mathbf{1} - A \text{ nicht invertierbar}\}.$$

Eine wichtige Eigenschaft des Spektrums ist, dass es nur von der Ähnlichkeitsklasse von  $A$  abhängt. In der Tat, sei  $B = MAM^{-1}$  für ein invertierbares  $M$ , dann gilt  $z\mathbf{1} - B = M(z\mathbf{1} - A)M^{-1}$  und somit ist die Invertierbarkeit von  $z\mathbf{1} - B$  gleichbedeutend zu der Invertierbarkeit von  $z\mathbf{1} - A$ . Also folgt

$$\text{Spec}(A) = \text{Spec}(B) = \text{Spec}(MAM^{-1}).$$

Wenn angewandt auf diagonalisierbare Matrizen suggeriert dies schon, dass das Spektrum Transformationseigenschaften hat. Die Annahmen in folgendem Satz sind ausreichend allgemein gehalten, damit sie auch auf interessante Funktionen angewandt werden können, die nicht auf ganz  $\mathbb{C}$  analytisch sind (z.B. die Wurzelfunktion).

**8.15 Satz** (Spektraler Abbildungssatz) Sei  $\Delta \subset \mathbb{C}$  offen und  $f : \Delta \rightarrow \mathbb{C}$  eine analytische Funktion gegeben durch eine konvergente Potenzreihe um 0. Ferner sei  $\text{Spec}(A) \subset \Delta$ . Dann gilt

$$f(\text{Spec}(A)) = \text{Spec}(f(A))$$

**Beweis.** Wir verwenden die Jordanzerlegung  $A = M(D + J)M^{-1}$  von  $A$ . Gemäß obiger Bemerkung reicht es also die Behauptung für  $A = D + J$  zu zeigen. Sei  $f(z) = \sum_{n \geq 0} f_n z^n$  mit  $f_n \in \mathbb{C}$ . Dann ist also

$$f(D + J) = \sum_{n \geq 0} f_n (D + J)^n = f(D) + S,$$

wobei  $S$  wieder eine nilpotente obere Dreiecksmatrix mit verschwindende Diagonale ist, denn  $[D, J] = 0$  und jede Potenz  $J^n$  eine obere Dreiecksmatrix mit verschwindende Diagonale ist. Also

$$f(D + J) = \begin{pmatrix} f(z'_1) & * & * \\ 0 & \ddots & * \\ 0 & 0 & f(z'_N) \end{pmatrix}, \quad D = \begin{pmatrix} z'_1 & & \\ & \ddots & \\ & & z'_N \end{pmatrix}.$$

Somit ist das charakteristische Polynom von  $f(D + J)$  leicht zu berechnen und die Behauptung folgt.  $\square$

## 9 Matrixzerlegungen

In diesem Kapitel werden lediglich Matrizen betrachtet und diese dann auf verschiedene Art und Weise zerlegt, d.h. als Matrixprodukt einfacherer Matrizen umgeschrieben. Alles überträgt sich dann auf abstrakte lineare Abbildungen, indem ihre darstellenden Matrizen betrachtet werden. Die Zerlegungen sind allesamt von Interesse sowohl in numerischen Anwendungen wie auch in theoretischen. Bei der Herleitung der Zerlegungen werden zudem noch eine Menge zusätzlicher Eigenschaften von Matrizen vorgestellt. Alles wird für komplexe Matrizen  $A \in \text{Mat}(N \times M, \mathbb{C})$  formuliert, die ja auch die reellen Matrizen enthalten. Der Bild- und Urbildraum der zu  $A$  zugehörigen linearen Abbildung sind  $\mathbb{C}^N$  und  $\mathbb{C}^M$ , und diese sind mit dem euklidischen Skalarprodukt versehen,

$$\langle v | w \rangle = \sum_{n=1}^N \bar{v}_n w_n, \quad v = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} \in \mathbb{C}^N, \quad w = \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix} \in \mathbb{C}^N.$$

Einige Matrixzerlegungen sind schon bekannt:

- Diagonalisierung  $A = M^{-1}DM$  für diagonalisierbares quadratisches  $A$ .
- Unitäre Diagonalisierung  $A = U^*DU$  für normales  $A$ .
- Jordanzerlegung  $A = M^{-1}(D + J)M$  für quadratisches  $A$

Auch die folgende Zerlegung ist im wesentlichen schon bekannt, denn es ist nichts anderes als eine Reformulierung der Gauss-Algorithmus. Es sei daran erinnert (Satz 5.4), dass eine Permutationsmatrix eine quadratische Matrix, die in jeder Zeile und jeder Spalte genau eine 1 und sonst nur 0 stehen hat.

**9.1 Satz** (LR-Zerlegung oder LU-Zerlegung oder Dreieckszerlegung) Sei  $A \in \text{Mat}(N \times M, \mathbb{C})$ . Dann existiert eine Permutationsmatrix  $P \in \text{Mat}(N \times N, \mathbb{C})$  und eine untere (linke, lower) Dreiecksmatrix  $L \in \text{Mat}(N \times N, \mathbb{C})$  und eine obere (rechte, upper) Dreiecksmatrix  $R \in \text{Mat}(N \times M, \mathbb{C})$ , so dass

$$PA = LR.$$

Die Diagonaleinträge von  $L$  können gleich 1 gewählt werden.

**Beweis.** Es wird der Gauss Algorithmus auf  $A$  wie im Beweis von Satz 4.20 angewandt, jedoch ohne Normierungsschritte. Wenn kein Vertauschen von Zeilen notwendig ist (d.h. alle Pivotelemente ungleich 0 sind), so werden also lediglich Vielfache vorheriger Zeilen dazuaddiert. Diese sind jeweils durch Multiplikation mit unteren Dreiecksmatrizen von links gegeben. Da das Produkt dieser unterer Dreiecksmatrizen wieder eine untere Dreiecksmatrix  $U$  mit 1 auf der Diagonale ist, folgt aus dem Gauss-Algorithmus, dass

$$UA = R,$$

wobei  $R$  eine obere Dreiecksmatrix ist. Nun ist  $U$  invertierbar und sein Inverses  $L = U^{-1}$  eine untere Dreiecksmatrix. Hieraus folgt die Zerlegung  $A = LR$ , die eindeutig ist. Wenn hingegen Zeilen vertauscht werden müssen, so kann das auch ganz zu Anfang geschehen und dies wird genau durch die Multiplikation eines geeigneten  $P$  von links erreicht.  $\square$

Es sei noch kurz der Algorithmus zur Bestimmung der LR-Zerlegung beschrieben in dem Fall wo keine Permutation notwendig ist (Kochrezept):

- Bilde  $(A|\mathbf{1}_N)$ .
- Wende den Gauss-Algorithmus auf  $(A|\mathbf{1}_N)$  an ohne zu normieren. Er liefert  $(R|\mathbf{1}_N - S)$  wobei  $S$  eine echte untere Dreiecksmatrix ist (insbesondere nilpotent).
- Dann ist  $L = (\mathbf{1}_N - S)^{-1} = \mathbf{1}_N + S + S^2 + \dots + S^{N-1}$ .

Eine weitere Faktorisierung, die mit Hilfe des Gauss-Algorithmus hergeleitet werden kann, ist die folgende Normalform. Es gibt von ihr Verallgemeinerungen auch für Matrizen mit Einträgen in beliebigen Ringen.

**9.2 Satz (Smith Normalform)** Sei  $A \in \text{Mat}(N \times M, \mathbb{C})$ . Dann existieren invertierbare Matrizen  $S \in \text{Mat}(N \times N, \mathbb{C})$  und  $T \in \text{Mat}(M \times M, \mathbb{C})$  sowie ein  $r \leq \min\{N, M\}$ , gleich dem Rang von  $A$ , mit

$$SAT = \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix}.$$

**Beweisskizze.** Dies ist im Wesentlichen der Rangsatz. Sei  $w_1, \dots, w_r$  eine Basis von  $\text{Ran}(A)$ . Dann sind die Urbilder  $v_1, \dots, v_r$  linear unabhängig in  $\mathbb{C}^M$  und können zu einer Basis ergänzt werden durch  $v_{r+1}, \dots, v_M$ . Dann ergänze noch  $w_1, \dots, w_r$  durch  $w_{r+1}, \dots, w_N$  zu einer Basis von  $\mathbb{C}^N$ . Dies ergibt  $S^{-1} = (w_1, \dots, w_N)$  und  $T = (v_1, \dots, v_N)$  die gewünschte Darstellung.  $\square$

Auch die nächste Zerlegung ist eine Umformulierung eines bekannten Algorithmus, nämlich des Gram-Schmidt'schen Orthonormalisierungsverfahrens angewandt auf die Zeilen der Matrix.

**9.3 Satz (QR Zerlegung)** Sei  $A \in \text{Mat}(N \times N, \mathbb{C})$  invertierbar. Dann existieren eine unitäre Matrix  $Q \in \text{Mat}(N \times N, \mathbb{C})$  und eine obere Dreiecksmatrix  $R \in \text{Mat}(N \times N, \mathbb{C})$  mit positiver Diagonale, so dass

$$A = QR.$$

Diese Zerlegung ist eindeutig.

**Beweis.** Es wird das Gram-Schmidt Orthonormierungsverfahren aus Satz 7.19 auf die Spaltenvektoren von  $A$  angewandt. Sei also  $A = (v_1, \dots, v_N)$  und dann werden  $u_1, \dots, u_N$  iterativ bestimmt durch

$$u_{n+1} = \frac{v_{n+1} - P_n v_{n+1}}{\|v_{n+1} - P_n v_{n+1}\|}, \quad P_n = \sum_{l=1}^n |u_l\rangle\langle u_l|.$$

Wenn wir dann setzen  $Q = (u_1, \dots, u_N)$ , so ist  $Q$  unitär. Weil der  $u_n$  lediglich Linearkombination von  $v_1, \dots, v_n$  ist, gilt zudem

$$Q = AO,$$

wobei  $O$  eine obere Dreiecksmatrix mit positiver Diagonale ist. Letzteres, weil der Diagonalkoeffizient einfach durch den positiven Normierungsfaktor gegeben ist. Nun ist  $R = O^{-1}$  eine obere Dreiecksmatrix mit positiver Diagonale (gegeben durch das Inverse der Diagonale von  $O$ ) und es gilt  $A = QR$ .  $\square$

Das Ziel ist es als nächstes die Polarzerlegung und die Singulärwertzerlegung vorzustellen. Hierfür sind jedoch eine lange Reihe an Vorbereitungen notwendig, die allerdings auch für sich genommen interessant sind. Sie betreffen Variationsprinzipien, positiv definite Matrizen und partielle Isometrien.

**9.4 Satz (Poincaré)** Sei  $A = A^* \in \text{Mat}(N \times N, \mathbb{C})$  und seien  $z'_n$  für  $n = 1, \dots, N$  die Eigenwerte aufgelistet mit ihrer Multiplizität und, so dass  $z'_{n+1} \leq z'_n$ . Sei  $U \subset \mathbb{C}^N$  ein beliebiger Unterraum der Dimension  $n$ . Dann gibt es normierte Vektoren  $v, w \in U$  mit

$$\langle v|Av \rangle \leq z'_n, \quad \langle w|Aw \rangle \geq z'_{N-n+1}.$$

**Beweis.** Sei  $V \subset \mathbb{C}^N$  der Unterraum, der von Eigenvektoren zu  $z'_n, \dots, z'_N$  aufgespannt wird. Dann gilt

$$\dim(U) + \dim(V) = n + (N - n + 1) = N + 1.$$

Somit erfüllt der Schnitt  $\dim(U \cap V) \geq 1$  nach Dimensionssatz (Satz 2.21). Sei  $v \in U \cap V$  normiert. Dann ist  $v = \sum_{l=n}^N \lambda_l u_l$ , wobei  $Au_l = z'_l u_l$  mit  $\langle u_k|u_l \rangle = \delta_{k,l}$  und  $\sum_{l=n}^N |\lambda_l|^2 = 1$ . Somit

$$\langle v|Av \rangle = \sum_{l,k=n}^N \bar{\lambda}_l \lambda_k \langle u_l|Au_k \rangle = \sum_{l,k=n}^N \bar{\lambda}_l \lambda_k \langle u_l|z'_k u_k \rangle = \sum_{l=n}^N |\lambda_l|^2 z'_l \leq \sum_{l=n}^N |\lambda_l|^2 z'_n = z'_n.$$

Dies zeigt die erste Ungleichung. Die zweite folgt, indem die erste auf  $-A$  angewandt wird, oder aber obiges Argument (gespiegelt) wiederholt wird.  $\square$

Hieraus folgen die ersten sogenannten Variationsprinzipien.

**9.5 Satz (Courant's Minimax-Prinzip und Fischer Prinzip)** Sei  $A = A^* \in \text{Mat}(N \times N, \mathbb{C})$  und seien  $z'_n$  für  $n = 1, \dots, N$  die Eigenwerte aufgelistet mit ihrer Multiplizität und, so dass  $z'_{n+1} \leq z'_n$ . Dann gilt

$$z'_n = \min_{\dim(U)=N-n+1} \max_{v \in U \setminus \{0\}} \frac{\langle v|Av \rangle}{\|v\|^2} = \max_{\dim(U)=n} \min_{v \in U \setminus \{0\}} \frac{\langle v|Av \rangle}{\|v\|^2},$$

wobei das Maximum bzw. Minimum berechnet wird, während  $U$  über alle Unterräume von  $\mathbb{C}^N$  mit angegebener Dimension variiert. Der Quotient auf der rechten Seite heißt auch der Rayleigh Koeffizient. Insbesondere gilt für den kleinsten und größten Eigenwert

$$z'_N = \min \text{Spec}(A) = \min_{v \neq 0} \frac{\langle v|Av \rangle}{\|v\|^2}, \quad z'_1 = \max \text{Spec}(A) = \max_{v \neq 0} \frac{\langle v|Av \rangle}{\|v\|^2}. \quad (9.1)$$

**Beweis.** Der Rayleigh Koeffizient ist skalierungsinvariant in  $v$  (d.h. unter der Abbildung  $v \mapsto \lambda v$ ). Somit reicht es zu zeigen, dass

$$z'_n = \max_{\dim(U)=n} \min_{v \in U, \|v\|=1} \langle v|Av \rangle = \min_{\dim(U)=N-n+1} \max_{v \in U, \|v\|=1} \langle v|Av \rangle.$$



Wir betrachten nun die erste Gleichung. Nach Satz 9.4 gibt es für jeden Unterraum  $U$  der Dimension  $n$  ein normiertes  $v \in U$  mit  $\langle v|Av \rangle \leq z'_n$ , also gilt für jedes  $U$ , dass  $\min_{v \in U, \|v\|=1} \langle v|Av \rangle \leq z'_n$ . Andererseits kann gewählt werden  $U = \text{span}\{u_1, \dots, u_n\}$ , wobei  $u_1, \dots, u_n$  ein Orthonormalsystem von Eigenvektoren mit  $Au_i = z'_i u_i$  ist. Für dieses spezielle  $U$  gilt  $\min_{v \in U, \|v\|=1} \langle v|Av \rangle = z'_n$ . Dies zeigt die erste Gleichung. Für die zweite kann analog vorgegangen werden, oder sie kann durch Vertauschen von  $A$  mit  $-A$  aus der ersten abgeleitet werden.  $\square$

**9.6 Definition** Eine Matrix  $A \in \text{Mat}(N \times N, \mathbb{C})$  heißt positiv semidefinit genau dann, wenn  $A = A^*$  und  $\text{Spec}(A) \subset [0, \infty)$ . Sie heißt positiv definit, falls zudem  $0 \notin \text{Spec}(A)$ . Analog definiert sind negativ definite und semidefinite Matrizen. Kurzschreibweisen hierfür sind  $A \geq 0$  und  $A > 0$ , bzw.  $A \leq 0$  und  $A < 0$ . Eine selbstadjungierte Matrix  $A$  heißt indefinit genau dann, wenn sie weder positiv semidefinit noch negativ semidefinit ist.

**9.7 Bemerkung** Es gilt  $A < 0$  genau dann, wenn  $-A > 0$ . In der Tat, das charakteristische Polynom erfüllt  $p_{-A}(z) = (-1)^N p_A(-z)$ , so dass die Nullstellen an 0 gespiegelt sind. Dies folgt z.B. auch aus dem spektralen Abbildungssatz angewandt mit der Funktion  $f(z) = -z$ .

**9.8 Satz** Sei  $A = A^* \in \text{Mat}(N \times N, \mathbb{C})$ . Dann

$$A \geq 0 \iff \langle v|Av \rangle \geq 0 \quad \forall v \in \mathbb{C}^N.$$

Analoges gilt für positive Definitheit, und negative Semidefinitheit und Definitheit.

**Beweis.** "  $\implies$  " Nach dem Spektralsatz für normale Matrizen gibt es eine unitäre Matrix  $U$ , so dass  $A = U^*DU$  und die Diagonalmatrix  $D$  hat nach Voraussetzung nur nicht-negative Einträge  $d_1, \dots, d_N$ . Nun folgt mit  $w = Uv$

$$\langle v|Av \rangle = \langle w|Dw \rangle = \sum_{n=1}^N d_n |w_n|^2 \geq 0, \quad w = \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix}.$$

Die Umkehrung "  $\impliedby$  " folgt aus (9.1), kann aber auch direkt wie folgt bewiesen werden. Nach dem Spektralsatz existiert eine Orthonormalbasis  $u_1, \dots, u_N$  von Eigenvektoren zu den Eigenwerte  $z'_1, \dots, z'_N$ . Dann ist nach Voraussetzung

$$z'_n = \langle u_n|Au_n \rangle \geq 0,$$

d.h. alle Eigenwerte von  $A$  sind nicht negativ.  $\square$

**9.9 Korollar** Sei  $A \in \text{Mat}(N \times N, \mathbb{C})$ . Dann gilt  $A^*A \geq 0$ . Falls  $A$  invertierbar ist, gilt  $A^*A > 0$ .

**Beweis.** Dies folgt direkt aus Satz 9.8, da  $\langle v|A^*Av \rangle = \langle Av|Av \rangle \geq 0$ , und echt größer als 0 falls  $A$  invertierbar ist.  $\square$

**9.10 Korollar** Sei  $A = A^* \in \text{Mat}(N \times N, \mathbb{C})$  und  $M \in \text{Mat}(N \times N, \mathbb{C})$  invertierbar. Dann gilt

$$A \geq 0 \iff M^*AM \geq 0.$$

Falls  $M$  nicht invertierbar ist, gilt nur die Implikation "  $\implies$  ". Analoges gilt für positive Definitheit, und negative Semidefinitheit und Definitheit.

**Beweis.** Dies folgt aus Satz 9.8, da  $\langle v|M^*AMv \rangle = \langle Mv|A(Mv) \rangle$  und  $M$  bijektiv ist.  $\square$

**9.11 Korollar** Die positiv definiten Matrizen bilden einen Kegel, d.h. wenn  $A > 0$ ,  $B > 0$  und  $\lambda > 0$ , dann ist auch  $A + \lambda B > 0$ .

**Beweis.** In der Tat, für jeden Vektor  $v$  gilt

$$\langle v|(A + \lambda B)v \rangle = \langle v|Av \rangle + \lambda \langle v|Bv \rangle > 0,$$

so dass die Aussage aus Satz 9.8 folgt. □

Nun sollen Kriterien vorgestellt werden, die es erlauben, positive Definitheit bei einer gegebenen Matrix zu überprüfen.

**9.12 Satz** Sei  $A = A^* \in \text{Mat}(N \times N, \mathbb{C})$  mit charakteristischem Polynom

$$p_A(z) = \det(A - z \mathbf{1}) = \sum_{n=0}^N p_n z^n.$$

Dann sind  $p_n \in \mathbb{R}$  und es gilt

$$(-1)^n p_n > 0 \text{ für } n = 0, \dots, N \iff A > 0.$$

**Beweis.** Die Aussage folgt aus *Descartes' Vorzeichenregel* für Polynome mit reellen Koeffizienten, die Positivität sämtlicher Wurzeln des Polynoms in Verbindung bringt mit alternierenden Vorzeichen, so wie angegeben. Der recht trickreiche Beweis hat nichts mit linearer Algebra zu tun und kann z.B. im Buch von Fischer nachgelesen werden [Fis]. □

**9.13 Satz** (Hurwitz' Hauptminorenkriterium) Sei  $A = A^* \in \text{Mat}(N \times N, \mathbb{C})$ . Für  $n = 1, \dots, N$  seien dann die folgenden Untermatrizen definiert:

$$A_n = (A_{k,l})_{1 \leq k,l \leq n} \in \text{Mat}(n \times n, \mathbb{C}).$$

Ihre Determinanten heißen Hauptminoren. Dann gilt

$$A > 0 \iff \det(A_n) > 0 \quad \forall n = 1, \dots, N.$$

**Beweis.** "⇒" Jeder Vektor  $v \in \mathbb{C}^n$  kann durch Nullen ergänzt werden zu einem Vektor  $\hat{v} \in \mathbb{C}^N$ . Dann gilt  $\langle v|A_n v \rangle = \langle \hat{v}|A \hat{v} \rangle$ . Nun ist Letzteres nach Voraussetzung immer positiv. Somit folgt aus Satz 9.8, dass  $A_n > 0$ . Die Determinante einer positiven Matrix ist als Produkt der positiven Eigenwerte aber positiv.

"⇐" Die Umkehrung wird durch Induktion über  $n$  gezeigt. Offensichtlich ist  $A_1 > 0$  (weil die Determinante ja der Diagonaleintrag ist). Es wird nun der Schritt von  $n - 1$  nach  $n$  überprüft. Sei also  $A_{n-1} > 0$ . Dann gibt es nach dem Spektralsatz eine Orthonormalbasis  $u_1, \dots, u_{n-1}$  von  $\mathbb{C}^{n-1}$  von Eigenvektoren zu  $A_{n-1}$ . Diese werden durch je eine Null ergänzt zu  $\hat{u}_1, \dots, \hat{u}_{n-1} \in \mathbb{C}^n$ . Diese Vektoren bilden ein Orthornormalsystem in  $\mathbb{C}^n$ , das durch den Standardbasisvektor  $e_n$  zu einer Basis ergänzt wird. Diesen letzten Basisvektor ersetzen wir nun durch

$$v = e_n - \sum_{k=1}^{n-1} \frac{1}{z_k} \langle \hat{u}_k | A_n e_n \rangle \hat{u}_k.$$

Dann setze  $M = (\hat{u}_1, \dots, \hat{u}_{n-1}, v) \in \text{Mat}(n \times n, \mathbb{C})$ , was invertierbar ist, da die Spaltenvektoren immer noch eine Basis bilden. Außerdem gilt

$$\begin{aligned} \langle \hat{u}_l | A_n v \rangle &= \langle \hat{u}_l | A_n e_n \rangle - \sum_{k=1}^{n-1} \frac{1}{z_k} \langle \hat{u}_k | A_n e_n \rangle \langle \hat{u}_l | A_n \hat{u}_k \rangle \\ &= \langle \hat{u}_l | A_n e_n \rangle - \sum_{k=1}^{n-1} \langle \hat{u}_k | A_n e_n \rangle \langle \hat{u}_l | \hat{u}_k \rangle = 0. \end{aligned}$$

Weil  $M^* A_n M$  selbstadjungiert ist, folgt deswegen

$$M^* A_n M = \begin{pmatrix} D & 0 \\ 0 & \langle v | A_n v \rangle \end{pmatrix}.$$

Nun ist nach Induktionsvoraussetzung  $D > 0$ . Außerdem besagt die Voraussetzung, dass die Determinante der linken Seite  $\det(M^* A_n M) = |\det(M)|^2 \det(A_n)$  positiv ist. Somit muss auch der letzte Diagonaleintrag  $\langle v | A_n v \rangle$  positiv sein. Es folgt, dass  $M^* A_n M > 0$  und nach Korollar 9.10 somit  $A_n > 0$ . Die Induktion besagt nun, dass  $A = A_N > 0$ .  $\square$

### 9.14 Beispiel Sei

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}.$$

Dann ist

$$p_A(z) = (2-z)^3 - (2-z) - (2-z) = -z^3 + 6z^2 - 10z + 4.$$

Also liegen alternierende Vorzeichen vor und somit  $A > 0$ . Auch das Hurwitz Kriterium kann angewandt werden, denn

$$\det(A_3) = \det(A) = 4, \quad \det(A_2) = \det \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} = 3, \quad \det(A_1) = 2,$$

so dass alle Hauptminoren positiv sind.  $\diamond$

**9.15 Definition** Der Betrag einer Matrix  $A \in \text{Mat}(N \times M, \mathbb{C})$  ist die positiv definite Matrix  $|A| = (A^* A)^{\frac{1}{2}} \in \text{Mat}(M \times M, \mathbb{C})$ , wobei die Wurzel durch das Spektralkalkül der positiven Matrix  $A^* A$  definiert ist, d.h. wenn

$$A^* A = U^* \begin{pmatrix} (\mu_1)^2 & & \\ & \ddots & \\ & & (\mu_M)^2 \end{pmatrix} U, \quad (9.2)$$

mit unitärem  $U \in \text{Mat}(M \times M, \mathbb{C})$ , dann

$$|A| = U^* \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_M \end{pmatrix} U.$$

Die Eigenwerte  $\mu_m$  von  $|A|$  heißen die Singulärwerte von  $A$  und werden geordnet durch

$$0 \leq \mu_M \leq \mu_{M-1} \leq \dots \leq \mu_1.$$

**9.16 Bemerkung** Selbst für eine quadratische Matrix  $A \in \text{Mat}(N \times N, \mathbb{C})$  gilt nicht immer  $|A| = |A^*|$ , es sei denn  $N = 1$  (für jede komplexe Zahl  $a$  gilt ja  $|\bar{a}| = |a|$ ). Es gilt  $|A| = |A^*|$  genau dann, wenn  $A$  normal ist. Satz 9.17 impliziert jedoch, dass zumindestens der Rang von  $|A|$  gleich dem von  $|A^*|$  ist. Eine weit stärkere Aussage ist, dass sogar die positiven Singulärwerte von  $A$  und  $A^*$  gleich sind.  $\diamond$

Folgende Tatsachen gelten auch in allgemeinerem Rahmen (siehe Satz 7.43).

**9.17 Satz** Sei  $A \in \text{Mat}(N \times M, \mathbb{C})$ . Dann gilt

$$\text{Ker}(A) = \text{Ran}(A^*)^\perp, \quad (9.3)$$

wobei das orthogonale Komplement in  $\mathbb{C}^M$  bez. des Standardskalarproduktes genommen wird. Zudem

$$\text{Ker}(A) = \text{Ker}(A^*A), \quad \text{Ran}(A) = \text{Ran}(AA^*).$$

Die Ränge von  $A$ ,  $A^*$ ,  $A^*A$ ,  $AA^*$ ,  $|A|$  und  $|A^*|$  sind allesamt gleich, d.h.

$$\dim(\text{Ran}(A)) = \dim(\text{Ran}(A^*)) = \dim(\text{Ran}(A^*A)) = \dim(\text{Ran}(AA^*)).$$

**Beweis.** Es gilt unter zweimaliger Verwendung von Lemma 7.27

$$\begin{aligned} v \in \text{Ker}(A) &\iff \langle w|Av \rangle = 0 \quad \forall w \in \mathbb{C}^N \\ &\iff \langle A^*w|v \rangle = 0 \quad \forall w \in \mathbb{C}^N \\ &\iff v \in \text{Ran}(A^*)^\perp, \end{aligned}$$

und die erste Behauptung ist gezeigt. Offensichtlich ist zunächst  $\text{Ker}(A) \subset \text{Ker}(A^*A)$ , aber für  $v \in \text{Ker}(A^*A)$  gilt

$$\|Av\|^2 = \langle Av|Av \rangle = \langle v|A^*Av \rangle = 0,$$

und somit auch  $Av = 0$ . Deswegen auch  $\text{Ker}(A^*A) \subset \text{Ker}(A)$ . Mit (9.3) folgt dann

$$\text{Ran}(A^*A) = \text{Ker}(A^*A)^\perp = \text{Ker}(A)^\perp = \text{Ran}(A^*),$$

also insbesondere auch  $\dim(\text{Ran}(A^*A)) = \dim(\text{Ran}(A^*))$ , und analog für  $AA^*$ . Aus (9.3) folgt

$$\begin{aligned} \dim(\text{Ran}(A^*)) &= M - \dim(\text{Ran}(A^*)^\perp) = M - \dim(\text{Ker}(A)) \\ &= M - (M - \dim(\text{Ran}(A))) = \dim(\text{Ran}(A)), \end{aligned}$$

wobei im vorletzten Schritt der Rangsatz für  $A$  verwandt wird.  $\square$

**9.18 Satz** (Singulärwertzerlegung) Sei  $A \in \text{Mat}(N \times M, \mathbb{C})$ . Dann gibt es unitäre Matrizen  $W \in \text{Mat}(N \times N, \mathbb{C})$  und  $U \in \text{Mat}(M \times M, \mathbb{C})$ , und eine Diagonalmatrix  $D \in \text{Mat}(N \times M, \mathbb{C})$  mit

$$A = WDU.$$

Auf der Diagonale von  $D$  stehen Singulärwerte von  $A$ .

**Beweis.** Um die Struktur der Konstruktion klar darzustellen, betrachten wir zunächst den Fall eines invertierbaren  $A$ , d.h. insbesondere  $N = M$ . Ausgangspunkt ist (9.2). Setze  $D = U|A|U^*$ . Dann ist  $D > 0$  und somit invertierbar. Deswegen ist es möglich,  $W = AU^*D^{-1}$  zu setzen, was ja zu  $A = WDU$  führt. Außerdem ist  $W$  in der Tat unitär, da

$$W^*W = (AU^*D^{-1})^*AU^*D^{-1} = D^{-1}UA^*AU^*D^{-1} = D^{-1}D^2D^{-1} = \mathbf{1}.$$

Nun betrachten wir den Fall  $M \geq N$ , der andere wird analog behandelt. Dann ist der Rang von  $A$  beschränkt durch  $\text{rk}(A) \leq N$  und somit gibt es höchstens  $N$  nicht verschwindende Singulärwerte, d.h.  $\mu_{N+1} = 0$  (vgl. Satz 7.43). Wir setzen

$$D = \begin{pmatrix} \mu_1 & & 0 & \dots & 0 \\ & \ddots & \vdots & & \vdots \\ & & \mu_N & 0 & \dots & 0 \end{pmatrix} \in \text{Mat}(N \times M, \mathbb{C}) .$$

Sei  $r$  so, dass  $\mu_r > 0$  und  $\mu_{r+1} = 0$  (d.h.  $r = \text{rk}(A)$  ist der Rang von  $A$ ). Wenn  $e_n \in \mathbb{C}^M$  die Standardbasisvektoren sind, dann gilt zunächst

$$AU^* e_n = 0, \quad n = r + 1, \dots, M . \quad (9.4)$$

Außerdem sind

$$w_n = \frac{1}{\mu_n} AU^* e_n, \quad n = 1, \dots, r, \quad (9.5)$$

wohldefiniert und erfüllen

$$\langle w_n | w_m \rangle = \frac{1}{\mu_n} \frac{1}{\mu_m} \langle AU^* e_n | AU^* e_m \rangle = \frac{1}{\mu_n} \frac{1}{\mu_m} \langle e_n | UA^* AU^* e_m \rangle = \delta_{n,m} .$$

Also bilden  $w_1, \dots, w_r$  ein Orthonormalsystem in  $\mathbb{C}^N$ , welches durch  $w_{r+1}, \dots, w_N$  zu einer Orthonormalbasis ergänzt wird. Dann ist  $W = (w_1, \dots, w_N)$  eine unitäre Matrix. Zusammen ergeben dann (9.4) und (9.5)

$$WD = AU^* ,$$

was eine Auswertung auf der Standardbasis zeigt. Dies impliziert die Zerlegung.  $\square$

**9.19 Bemerkung** Der obige Beweis kann durchaus auch als Algorithmus (Kochrezept) zur Bestimmung der Singulärwertzerlegung verstanden werden:

- Berechne  $A^*A$  führe eine unitäre Diagonalisierung  $A^*A = U^* \widehat{D} U$  durch.
- Bilde  $D$  (von gleicher Größe wie  $A$ ) aus  $\widehat{D}$  durch hinzufügen von Nullzeilen oder Nullspalten.
- Berechne die Bildvektoren  $w_n = \frac{1}{\mu_n} AU^* e_n$  für  $n = 1, \dots, r = \text{rk}(A)$ .
- Vervollständige  $w_1, \dots, w_r$  zu einer Orthonormalbasis (z.B. durch Gram-Schmidt) und bilde  $W$ .

Alternativ kann auch erst die Singulärwertzerlegung von  $A^*$  berechnet werden und daraus die von  $A$  abgelesen werden (durch Adjungieren). Dies kann weniger Rechenaufwand bedeuten, wenn  $M > N$  so dass  $AA^*$  kleiner als  $A^*A$  ist.  $\diamond$

**9.20 Satz** (Polarzerlegung) Sei  $A \in \text{Mat}(N \times N, \mathbb{C})$ . Dann existiert eine unitäre Matrix  $V \in \text{Mat}(N \times N, \mathbb{C})$  mit

$$A = V |A| .$$

**Beweis.** Aus der Singulärwertzerlegung folgt

$$A = WDU = WUU^*DU = WU|A| ,$$

wobei  $W$  und  $U$  unitär sind. Also ist auch  $V = WU$  unitär.

Im Fall  $\text{Ker}(A) = \{0\}$  kann auch ein kurzer Beweis ohne Singulärwertzerlegung angegeben werden. Dann ist nämlich  $|A|$  invertierbar und somit kann man  $V = A|A|^{-1}$  setzen. In der Tat folgt dann aus  $A^*A = |A|^*|A|$  auch  $(|A|^*)^{-1}A^*A|A|^{-1} = \mathbf{1}$ , was gleichbedeutend mit  $V^*V = \mathbf{1}$ .  $\square$

Es gibt eine weitere Version der Polarzerlegung, die anstelle der Unitären eine eindeutigen sogenannte partielle Isometrie verwendet.

**9.21 Definition** Eine Matrix  $V \in \text{Mat}(N \times M, \mathbb{C})$  heißt partielle Isometrie, falls  $V^*V$  eine Projektion im  $\mathbb{C}^M$  ist.

**9.22 Satz** Eine Matrix  $V \in \text{Mat}(N \times M, \mathbb{C})$  ist genau dann eine partielle Isometrie, wenn es ein  $K$  gibt und Orthonormalsysteme  $u_1, \dots, u_K$  in  $\mathbb{C}^M$  und  $w_1, \dots, w_K$  in  $\mathbb{C}^N$  dergestalt, dass

$$V = \sum_{k=1}^K |w_k\rangle\langle u_k|.$$

Hierbei spannen  $u_1, \dots, u_K$  den Unterraum  $\text{Ran}(V^*) \subset \mathbb{C}^M$  auf, und  $w_1, \dots, w_K$  den Unterraum  $\text{Ran}(V) \subset \mathbb{C}^N$ . Außerdem ist  $V$  eine partielle Isometrie genau dann, wenn auch  $V^*$  eine partielle Isometrie ist.

**Beweis.** Da  $V^*V$  eine Projektion auf den Unterraum  $\text{Ran}(V^*V) = \text{Ran}(V^*)$  ist, gilt nach Satz 7.40

$$V^*V = \sum_{k=1}^K |u_k\rangle\langle u_k|.$$

Nun setze  $w_k = Vu_k$ . Diese Vektoren bilden in der Tat ein Orthonormalsystem, denn

$$\langle w_k | w_l \rangle = \langle Vu_k | Vu_l \rangle = \langle u_k | V^*V u_l \rangle = \delta_{k,l}.$$

Außerdem gilt

$$V = \sum_{k=1}^K V|u_k\rangle\langle u_k| = \sum_{k=1}^K |w_k\rangle\langle u_k|,$$

denn jede Matrix ist nach Satz 4.26 durch die Wirkung auf einer Basis festgelegt (und die ist hier  $u_1, \dots, u_K$  ergänzt durch eine Basis von  $\text{Ker}(A)$ ). Somit ist gezeigt, dass  $V$  die angegebene Darstellungsformel hat. Umgekehrt, wenn  $V = \sum_{k=1}^K |w_k\rangle\langle u_k|$ , so kann direkt nachgerechnet werden, dass  $V^*V$  eine Projektion ist, d.h.  $V$  eine partielle Isometrie.

Somit sind die ersten Aussagen verifiziert. Nun zur letzten Äquivalenz. Wenn  $V$  wie oben ist, dann ist

$$V^* = \sum_{k=1}^K |u_k\rangle\langle w_k| \in \text{Mat}(M \times N, \mathbb{C}).$$

Diese Abbildung ist nach obiger Charakterisierung aber wieder eine partielle Isometrie. □

**9.23 Bemerkung** Es sei daran erinnert (Bemerkung 7.41), dass für jede Projektion  $P$  und jeden Vektor  $v \in \text{Ran}(P)$  auch gilt  $Pv = v$  (weil  $v = Pw = P^2w = PPw = Pv$ ). Somit gilt

$$\|Vv\|^2 = \langle Vv | Vv \rangle = \langle v | V^*Vv \rangle = \langle v | v \rangle = \|v\|^2, \quad v \in \text{Ran}(V^*V) = \text{Ran}(V^*) = \text{Ker}(V)^\perp.$$

Für Vektoren  $w \in \text{Ker}(V)$  gilt  $\|Vw\| = 0$ . Aus diesen Gründen wird  $V$  eine partielle Isometrie, weil eben nur für einige Vektoren die Länge erhalten bleibt. Zusammenfassend ist die partielle Isometrie spezifiziert durch zwei Zerlegungen:

$$\begin{aligned} \mathbb{C}^M &= \text{Ker}(V) \oplus \text{Ker}(V)^\perp = \text{Ker}(V) \oplus \text{Ran}(V^*), \\ \mathbb{C}^N &= \text{Ker}(V^*) \oplus \text{Ker}(V^*)^\perp = \text{Ker}(V^*) \oplus \text{Ran}(V), \end{aligned}$$

zusammen mit einer isometrische Bijektion von  $\text{Ran}(V^*)$  auf  $\text{Ran}(V)$ . Es gibt im allgemeinen viele solche Abbildungen, d.h. die partielle Isometrie ist nicht eindeutig festgelegt durch die Vorgabe von zwei Zerlegungen und  $\mathbb{C}^N$  und  $\mathbb{C}^M$  (mit gleichdimensionalen zweiten Summanden). ◇

**9.24 Satz** (Eindeutige Polarzerlegung mit partieller Isometrie) Sei  $A \in \text{Mat}(N \times M, \mathbb{C})$ . Dann existiert eine eindeutige partielle Isometrie  $V \in \text{Mat}(N \times M, \mathbb{C})$  mit

$$A = V|A|, \quad \text{Ker}(V) = \text{Ker}(A).$$

Es gilt: (i)  $V^*V|A| = |A|$  (ii)  $V^*A = |A|$  (iii)  $VV^*A = A$

**Beweis.** Die Konstruktion ist wie in Satz 9.20 besonders einfach im Fall  $\text{Ker}(A) = \{0\}$ , denn dann ist  $|A|$  invertierbar. Dann setzt man  $V = A|A|^{-1}$ . In der Tat folgt dann aus  $A^*A = |A|^*|A|$  auch  $(|A|^*)^{-1}A^*A|A|^{-1} = \mathbf{1}$ , was gleichbedeutend mit  $V^*V = \mathbf{1}$  ist.

Im allgemeinen Fall beginnen wir mit der Singulärwertzerlegung. Hierin ist  $D \in \text{Mat}(N \times M, \mathbb{C})$  mit den Singulärwerten als Diagonaleinträgen. Wenn  $r$  der Rang von  $A$  ist, so kann dies auch zerlegt werden in

$$D = \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} \hat{D},$$

wobei

$$\hat{D} = \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_M \end{pmatrix} \in \text{Mat}(M \times M, \mathbb{C}).$$

Somit wird die Singulärwertzerlegung

$$A = W \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} \hat{D} U = W \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} U U^* \hat{D} U = W \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} U |A|,$$

wo im letzten Schritt Definition 9.15 verwandt wurde. Nun setze

$$V = W \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} U.$$

Es gilt

$$V^*V = U^* \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} U,$$

was eine Projektion ist, denn  $V^*V$  ist selbstadjungiert und  $V^*VV^*V = V^*V$ . Also ist  $V$  in der Tat eine partielle Isometrie mit

$$\text{Ker}(V) = \text{Ker} \left( \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} U \right) = \text{Ker}(DU) = \text{Ker}(U^*D^2U) = \text{Ker}(A^*A) = \text{Ker}(A).$$

Dies beendet den Beweis der Existenz der Polarzerlegung, ausgehend von der Singulärwertzerlegung.

Es sei noch kurz ein direktes alternatives Argument für die Existenz skizziert. Zunächst gilt für alle  $v \in \mathbb{C}^N$

$$\| |A|v \|^2 = \langle |A|v \mid |A|v \rangle = \langle v \mid |A|^2 v \rangle = \langle v \mid A^*A v \rangle = \|Av\|^2. \quad (9.6)$$

Also ist  $\text{Ker}(|A|) = \text{Ker}(A)$ , was auch schon aus Satz 9.17 folgt. Nun definieren wir  $V : \text{Ran}(|A|) \rightarrow \text{Ran}(A)$  als lineare Abbildung durch

$$V|A|v = Av.$$

Nach (9.6) ist  $V$  wohldefiniert, d.h.  $|A|v = |A|w$  impliziert  $|A|(v-w) = 0$  und somit  $A(v-w) = 0$ , d.h.  $Av = Aw$ . Zudem ist  $V$  nach (9.6) isometrisch. Um  $V$  zu einer auf ganz  $\mathbb{C}^M$  definierten linearen Abbildung zu

machen, wird nun  $V$  auf  $\text{Ran}(|A|)^\perp = \text{Ker}(|A|^*) = \text{Ker}(|A|)$  gleich Null gesetzt, d.h.  $\text{Ker}(V) = \text{Ker}(|A|) = \text{ker}(A)$ . Nach Bemerkung 9.23 ist  $V$  dann eine partielle Isometrie.

Zur Eindeutigkeit: Sei  $\hat{V}$  eine weitere partielle Isometrie mit

$$\hat{V}|A| = V|A| = A, \quad \text{Ker}(\hat{V}) = \text{Ker}(A) = \text{Ker}(V) = \text{Ker}(|A|).$$

Aus Ersterem folgt  $\hat{V}v = Vv$  für alle  $v \in \text{Ran}(|A|) = \text{Ker}(|A|)^\perp$ , was zusammen mit Letzterem  $\hat{V} = V$  impliziert.

Zuletzt zu den Identitäten: (i)  $V^*V$  ist die Projektion auf  $\text{Ker}(V)^\perp = \text{Ker}(A)^\perp = \text{Ker}(|A|)^\perp = \text{Ran}(|A|)^\perp$ . Also folgt  $V^*V|A| = |A|$ . (ii) Mit (i) folgt  $V^*A = V^*V|A| = |A|$ . (iii)  $VV^*A = VV^*V|A| = V|A| = A$ , wieder nach (i).  $\square$

**9.25 Beispiel** Es soll der Betrag sowie die Singulärwertzerlegung als auch die eindeutige Polarzerlegung (mit partieller Isometrie wie in Satz 9.24) von

$$A = \begin{pmatrix} 1 & -i \\ 0 & 1 \\ -i & 0 \end{pmatrix}$$

bestimmt werden. Zunächst ist

$$A^*A = \begin{pmatrix} 1 & 0 & i \\ i & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -i \\ 0 & 1 \\ -i & 0 \end{pmatrix} = \begin{pmatrix} 2 & -i \\ i & 2 \end{pmatrix}.$$

Das charakteristische Polynom ist  $p_{A^*A}(z) = z^2 - 4z + 3 = (z - 3)(z - 1)$ . Somit sind die Eigenwerte  $z_1 = 3$  und  $z_2 = 1$  mit orthonormierten Eigenvektoren

$$u_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}, \quad u_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}.$$

Also setzen wir

$$U^* = (u_1, u_2) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}.$$

Dann ist

$$A^*A = U^* \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} U,$$

und somit

$$|A| = U^* \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{pmatrix} U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 + \sqrt{3} & i(1 - \sqrt{3}) \\ -i(1 - \sqrt{3}) & 1 + \sqrt{3} \end{pmatrix}.$$

In der Singulärwertzerlegung  $A = WDU$  sind nun bekannt  $U$  und

$$D = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Gemäß (9.5) werden nun berechnet

$$w_1 = \frac{1}{\sqrt{3}} AU^* e_1 = \frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ i \\ -i \end{pmatrix}, \quad w_2 = AU^* e_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -i \\ -i \end{pmatrix}.$$



Diese Vektoren werden ergänzt zu einer Orthonormalbasis, z.B. durch Anwenden des Gram-Schmidt Verfahrens angewandt auf den dritten Vektor  $e_1$ . Dies führt zu

$$w_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -i \\ i \end{pmatrix}.$$

Dann ist  $W = (w_1, w_2, w_3)$ . Zusammenfassend

$$A = \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{i}{\sqrt{6}} & \frac{-i}{\sqrt{2}} & \frac{-i}{\sqrt{3}} \\ \frac{-i}{\sqrt{6}} & \frac{-i}{\sqrt{2}} & \frac{-i}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}.$$

Für die Polarzerlegung  $A = V|A|$  können wir nun  $V = A|A|^{-1}$  verwenden, denn das Inverse  $|A|^{-1}$  existiert und ist

$$|A|^{-1} = U^* \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 \\ 0 & 1 \end{pmatrix} U = \frac{1}{2} \begin{pmatrix} 1 + \frac{1}{\sqrt{3}} & i(1 - \frac{1}{\sqrt{3}}) \\ -i(1 - \frac{1}{\sqrt{3}}) & 1 + \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Also

$$V = A|A|^{-1} = \begin{pmatrix} 1 & -i \\ 0 & 1 \\ -i & 0 \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 + \frac{1}{\sqrt{3}} & i(1 - \frac{1}{\sqrt{3}}) \\ -i(1 - \frac{1}{\sqrt{3}}) & 1 + \frac{1}{\sqrt{3}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{i}{\sqrt{3}} \\ -\frac{i}{2}(1 - \frac{1}{\sqrt{3}}) & \frac{1}{2}(1 + \frac{1}{\sqrt{3}}) \\ -\frac{i}{2}(1 + \frac{1}{\sqrt{3}}) & \frac{1}{2}(1 - \frac{1}{\sqrt{3}}) \end{pmatrix}.$$

Alternativ kann  $V = W \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} U$  verwandt werden. Somit sind alle Matrizen berechnet.  $\diamond$

Nun soll noch eine Anwendung der Singulärwertzerlegung angegeben werden, und zwar Gauss' Methode der kleinsten Fehlerquadrate und die damit verbundene Pseudoinverse (oder Moore-Penrose Inverse, 1920, 1955). Die Ausgangsfragestellung ist die folgende. Ein lineares Gleichungssystem  $Ax = b$  mit  $A \in \text{Mat}(N \times M, \mathbb{C})$  habe keine Lösung. Was ist dann die beste Approximation einer Lösung, d.h. ein  $x \in \mathbb{C}^M$ , für welches

$$\|Ax - b\|^2 = \sum_{n=1}^N \left| \sum_{m=1}^M A_{n,m} x_m - b_n \right|^2$$

minimal ist? Dieses Minimum muss nicht unbedingt eindeutig sein, aber es gibt ein eindeutiges kürzestes  $\hat{x} \in \mathbb{C}^M$  unter diesen Minima, für welches zudem  $\|x\|$  minimal ist. Geometrisch betrachtet ist  $b$  nicht im Bild  $\text{Ran}(A)$ , aber man kann nach dem Punkt in  $\text{Ran}(A)$  suchen, der den Abstand zu  $b$  minimiert. Dieser ist durch die orthogonale Projektionen  $Pb$  von  $b$  auf  $\text{Ran}(A)$  gegeben. Nun hat  $Pb$  aber evtl. viele Urbildpunkte, und von denen wird der kürzeste ausgewählt. Was wir im Folgenden zeigen werden ist, dass  $\hat{x}$  durch das Pseudoinverse  $A^+$  berechnet werden kann durch  $\hat{x} = A^+b$ . Zunächst wird diese Pseudoinverse eingeführt und dann die Lösungseigenschaft nachgewiesen.

**9.26 Satz** (Pseudoinverse  $A^+$  nach Moore-Penrose) Sei  $A \in \text{Mat}(N \times M, \mathbb{C})$ . Dann gibt es genau ein  $A^+ \in \text{Mat}(M \times N, \mathbb{C})$  mit

$$A = AA^+A, \quad A^+ = A^+AA^+, \quad (A^+A)^* = A^+A, \quad (AA^+)^* = AA^+. \quad (9.7)$$

Falls  $A$  reell ist, so ist auch  $A^+$  reell.

**Beweis.** Wir beginnen mit einer Singulärwertzerlegung

$$A = WDU, \quad D = \begin{pmatrix} \mu_1 & & & 0 & \dots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & \mu_r & \vdots & & \vdots \\ & & & 0 & & \vdots \\ & & & & \ddots & \vdots \\ & & & & & 0 & 0 & \dots & 0 \end{pmatrix} \in \text{Mat}(N \times M, \mathbb{C}),$$

wobei  $r = \text{rk}(A)$  der Rang von  $A$  ist, d.h.  $\mu_r > 0$  und  $\mu_{r+1} = 0$ , und setzen

$$A^+ = U^*D^+W^*, \quad D^+ = \begin{pmatrix} (\mu_1)^{-1} & & & 0 & \dots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & (\mu_r)^{-1} & \vdots & & \vdots \\ & & & 0 & & \vdots \\ & & & & \ddots & \vdots \\ & & & & & 0 & 0 & \dots & 0 \end{pmatrix} \in \text{Mat}(M \times N, \mathbb{C}).$$

Dies ist wieder ein Operator gegeben in einer Singulärwertzerlegung. Beachten Sie, dass auch gilt

$$D^+ = (D^*D)^{-1}D^*.$$

Nun werden die 4 Identitäten (9.7) nachgerechnet. Zunächst sind

$$AA^+ = WDUU^*D^+W^* = WDD^+W^* = W \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} W^*, \quad A^+A = U^* \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} U,$$

beide selbstadjungiert, und

$$AA^+A = W \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} W^*WDU = WDU = A,$$

sowie

$$A^+AA^+ = U^* \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} UU^*D^+W^* = U^*D^+W^* = A^+.$$

Somit zeigt die Konstruktion die Existenz von  $A^+$ . Nun soll noch die Eindeutigkeit nachgewiesen werden. Sei  $B \in \text{Mat}(M \times N, \mathbb{C})$  ein weiteres Pseudoinverses. Dann ist

$$\begin{aligned} (A^+A - BA)^*(A^+A - BA) &= (A^+A - BA)(A^+A - BA) \\ &= A^+AA^+A - A^+ABA - BAA^+A + BABA \\ &= A^+A - A^+A - BA + BA = 0. \end{aligned}$$

Also  $A^+A - BA = 0$  (wenn  $C^*C = 0$ , dann  $C = 0$  nach Satz 9.17) und somit  $A^+A = BA$ . Analog zeigt man

$$(AA^+ - AB)^*(AA^+ - AB) = 0,$$

so dass auch  $AA^+ = AB$ . Also, unter Verwendung der anderen Eigenschaften von Pseudoinversen,

$$B = BAB = A^+AB = A^+AA^+ = A^+,$$

und der Beweis ist beendet. □

**9.27 Bemerkung** Der obige Beweis liefert auch einen Algorithmus (Kochrezept) zur Berechnung des Pseudoinversen.

- Berechne die Singulärwertzerlegung  $A = WDU$  wie oben angegeben, oder  $A^* = U^*D^*W^*$ .
- Bilde  $D^+ = (D^*D)^{-1}D^*$  und  $A^+ = U^*D^+W^*$ .

◇

**9.28 Beispiel** Oben wurde folgende Singulärwertzerlegung berechnet:

$$A = \begin{pmatrix} 1 & -i \\ 0 & 1 \\ -i & 0 \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{i}{\sqrt{6}} & \frac{-i}{\sqrt{2}} & \frac{-i}{\sqrt{3}} \\ \frac{-i}{\sqrt{6}} & \frac{-i}{\sqrt{2}} & \frac{-i}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}.$$

Nach der Konstruktion im Satz ist das Pseudoinverse also durch

$$A^+ = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{6}} & \frac{-i}{\sqrt{6}} & \frac{i}{\sqrt{6}} \\ 0 & \frac{i}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{i}{\sqrt{3}} & \frac{i}{\sqrt{3}} \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 & i & 2i \\ i & 2 & 1 \end{pmatrix},$$

gegeben.

◇

Als Nächstes sollen Eigenschaften des Pseudoinversen verifiziert werden.

**9.29 Satz** (Eigenschaften der Pseudoinversen) Sei  $A^+$  Pseudoinverses zu  $A \in \text{Mat}(N \times M, \mathbb{C})$ . Dann:

- (i)  $(A^+)^+ = A$
- (ii)  $(A^*)^+ = (A^+)^*$
- (iii)  $\text{rk}(A) = \text{rk}(A^+)$
- (iv)  $A^+ = A^{-1}$ , falls  $A$  invertierbar ist.
- (v) Falls  $\text{rk}(A) = M$  (maximaler Spaltenrang),  $A^+ = (A^*A)^{-1}A^*$ .
- (vi) Für unitäre  $\tilde{U}$  und  $\tilde{W}$  richtiger Größe gilt  $(\tilde{U}A\tilde{W})^+ = \tilde{W}^*A^+\tilde{U}^*$ .

**Beweis.** Es sei daran erinnert, dass  $A$  und  $A^+$  jeweils durch eine Singulärwertzerlegung gegeben sind:

$$A = WDU, \quad A^+ = U^*D^+W^*,$$

wobei  $D$  und  $D^+$  wie oben und insbesondere von gleichem Rang sind. Hieraus, kombiniert mit der Eindeutigkeit von  $A^+$ , folgen sofort alle Behauptungen, bis auf (v). Nun also zu Letzterem. Aus  $r = \text{rk}(A) = M$  folgt  $\text{Ran}(A^*) = \mathbb{C}^M$  und somit ist  $A^*A$  invertierbar und gegeben durch  $A^*A = U^*D^*DU$ . Also  $(A^*A)^{-1}A^* = U^*(D^*D)^{-1}D^*W^* = A^+$ , weil nämlich  $D^+ = (D^*D)^{-1}D^*$ . □

**9.30 Satz** (Gauss' Methode der kleinsten Quadrate) Sei  $A \in \text{Mat}(N \times M, \mathbb{C})$  und  $b \in \mathbb{C}^N$ . Dann ist  $\hat{x} = A^+b$  die eindeutige Lösung des folgenden Optimierungsproblems:

$$\|A\hat{x} - b\| = \min_{x \in \mathbb{C}^M} \|Ax - b\|, \quad \|\hat{x}\| = \min \{ \|x\| \mid x \in \mathbb{C}^M \text{ mit } \|Ax - b\| = \|A\hat{x} - b\| \}.$$

**Beweis.** Sei  $\hat{x} = A^+b$  und  $x \in \mathbb{C}^M$  beliebig. Dann gilt unter Verwendung der Eigenschaften der Pseudoinversen

$$\begin{aligned} \langle A\hat{x} - b | A(x - \hat{x}) \rangle &= \langle (AA^+ - \mathbf{1})b | A(x - \hat{x}) \rangle = \langle b | (AA^+ - \mathbf{1})^* A(x - \hat{x}) \rangle \\ &= \langle b | (AA^+ - \mathbf{1})A(x - \hat{x}) \rangle = \langle b | (A - A)(x - \hat{x}) \rangle = 0. \end{aligned}$$

Also ist  $A\hat{x} - b$  orthogonal zu  $A(x - \hat{x})$ . Somit impliziert der Satz von Pythagoras

$$\|Ax - b\|^2 = \|A(x - \hat{x}) + (A\hat{x} - b)\|^2 = \|A(x - \hat{x})\|^2 + \|(A\hat{x} - b)\|^2 \geq \|(A\hat{x} - b)\|^2.$$

Hierbei ist die letzte Ungleichung eine Gleichheit genau dann, wenn  $A(x - \hat{x}) = 0$  gilt, d.h.  $x - \hat{x} \in \text{Ker}(A)$ . Andererseits gilt  $\hat{x} \in \text{Ker}(A)^\perp$ , weil nämlich für jedes  $y \in \text{Ker}(A)$

$$\langle \hat{x} | y \rangle = \langle A^+b | y \rangle = \langle A^+AA^+b | y \rangle = \langle A^+b | (A^+A)^*y \rangle = \langle A^+b | A^+Ay \rangle = 0.$$

Deswegen gilt für  $x$  mit  $x - \hat{x} \in \text{Ker}(A)$  (d.h.  $x$  minimiert  $\|Ax - b\|^2$ ) auch

$$\|x\|^2 = \|x - \hat{x} + \hat{x}\|^2 = \|x - \hat{x}\|^2 + \|\hat{x}\|^2 \geq \|\hat{x}\|^2,$$

wobei wieder der Satz von Pythagoras angewandt wurde. Dies zeigt, dass  $\hat{x}$  der gesuchte Minimierer ist.  $\square$

**9.31 Bemerkung** Die Gleichung  $\|A\hat{x} - b\| = \min_{x \in \mathbb{C}^M} \|Ax - b\|$  besagt geometrisch, dass  $A\hat{x}$  genau die orthogonale Projektion von  $b$  auf den Unterraum  $\text{Ran}(A)$  ist, wie bei der Problemeinführung erläutert (vgl. auch eine der Übungsaufgaben zu Projektionen).

**9.32 Beispiel** Sei

$$A = \begin{pmatrix} 1 & -i \\ 0 & 1 \\ -i & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

Oben wurde schon  $A^+$  berechnet. Nun ist die im obigen Sinne beste approximative Lösung von  $Ax = b$  gegeben durch

$$\hat{x} = A^+b = \frac{1}{3} \begin{pmatrix} 1 & i & 2i \\ i & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 - i \\ 1 + i \end{pmatrix}.$$

Beachte hierbei, dass  $b \notin \text{Ran}(A)$ , weil  $\det(A, b) \neq 0$ , und somit  $\hat{x}$  keine Lösung von  $Ax = b$  ist.  $\diamond$

Als Praxisbeispiel sei die Kurvenanpassung von gegebenen Daten mit Hilfe der Methode der kleinsten Fehlerquadrate vorgestellt. Hierbei hat man einen Satz von reellen Messdaten  $(x_n, y_n)_{n=1, \dots, N}$  und eine Familie  $(f_m)_{m=1, \dots, M}$  von Funktionen  $f_m : \mathbb{R} \rightarrow \mathbb{R}$ , deren Spann zur Beschreibung der Daten in Frage kommen. Meist ist hierbei  $M \ll N$ . Nun soll die beste Funktion  $f = \sum_{m=1}^M v_m f_m$  mit  $v_m \in \mathbb{R}$  bestimmt werden, so dass

$$\sum_{n=1}^N |f(x_n) - y_n|^2,$$

minimiert wird, und gleichzeitig auch  $\sum_{m=1}^M |v_m|^2$ . Ein typisches Beispiel sind die Funktionen  $f_m(x) = x^{m-1}$ , denn dann wird nach dem besten Polynom vom Grade  $M - 1$  gefragt, das die Messdaten im Sinne

kleiner Fehlerquadrate am besten approximiert (siehe Beispiel unten). Nun kann dieses Problem wie folgt umformuliert werden. Setze

$$A = (f_m(x_n))_{n=1,\dots,N, m=1,\dots,M} \in \text{Mat}(N \times M, \mathbb{R}), \quad v = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \in \mathbb{R}^M.$$

Dann ist  $f(x_n) = (Av)_n$  und gesucht ist genau der Vektor  $\hat{v} \in \mathbb{R}^M$ , so dass  $\|Av - y\|$  und  $\|v\|$  beide minimiert werden. Die Lösung dieses Problems ist dann gegeben durch

$$\hat{v} = A^+ y, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix}.$$

Bei all dem ist zu beachten, dass  $A$  nur von den Fußpunkten  $x_n$  und Funktionen  $f_m$  abhängt. Wenn also für andere  $y_1, \dots, y_N$  eine weitere beste Approximation gesucht wird, so kann das gleiche Pseudoinverse verwendet werden.

**9.33 Beispiel** Gegeben seien Messdaten  $(x_1, y_1) = (1, 1)$ ,  $(x_2, y_2) = (2, 4)$  und  $(x_3, y_3) = (0, 0)$ . Diese Daten sollen durch eine Gerade  $f(x) = v_0 + v_1x$  im Sinne der kleinsten Fehler optimal approximiert werden. Dann sind also  $f_0(x) = 1$  und  $f_1(x) = x$  sowie  $M = 2$  und  $N = 3$ . Deswegen

$$A = \begin{pmatrix} f_0(x_1) & f_1(x_1) \\ f_0(x_2) & f_1(x_2) \\ f_0(x_3) & f_1(x_3) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 0 \end{pmatrix}.$$

Die Pseudoinverse ist (Wolfram mathematica, oder zu Fuß berechnen)

$$A^+ = \begin{pmatrix} \frac{1}{3} & -\frac{1}{6} & \frac{5}{6} \\ 0 & \frac{1}{2} & -\frac{1}{2} \end{pmatrix},$$

und somit

$$\hat{v} = A^+ y = \begin{pmatrix} \frac{1}{3} & -\frac{1}{6} & \frac{5}{6} \\ 0 & \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} \\ 2 \end{pmatrix}.$$

Also ist die beste approximierende Gerade  $f(x) = -\frac{1}{3} + 2x$ , d.h. die Verbindungsgerade von  $(0, 0)$  nach  $(2, 4)$  sollte wegen  $(1, 1)$  etwas nach unten versetzt werden, genau um  $\frac{1}{3}$ .  $\diamond$

## 10 Sesquilinearformen, hermitische Formen und Quadriken

In Definition 5.14 wurden schon Bilinearformen auf einem beliebigen Vektorraum eingeführt. Eine Abbildung  $B : V \times V \rightarrow \mathbb{K}$  heißt eine Bilinearform, falls  $B$  linear in beiden Argumenten ist, d.h.

$$B(v + \lambda w, v' + \lambda' w') = B(v, v') + \lambda B(w, v') + \lambda' B(v, w') + \lambda \lambda' B(w, w').$$

Ein solche Bilinearform ist symmetrisch, wenn  $B(v, w) = B(w, v)$  für alle  $v, w \in V$ , und sie ist antisymmetrisch, wenn  $B(v, w) = -B(w, v)$  für alle  $v, w \in V$ . Beispiel 5.15 stellt Bilinearformen  $B : \mathbb{C}^N \times \mathbb{C}^N \rightarrow \mathbb{C}$  folgender Gestalt vor:

$$B(v, w) = \sum_{m,n=1}^N v_m A_{m,n} w_n, \quad v = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix}, \quad w = \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix}, \quad (10.1)$$

wobei  $A \in \text{Mat}(N \times N, \mathbb{K})$ . Diese Form ist symmetrisch genau dann, wenn  $A^t = A$  symmetrisch ist, und antisymmetrisch genau dann, wenn  $A^t = -A$  antisymmetrisch ist. Tatsächlich ist es möglich zu zeigen, dass jede Bilinearform auf einem endlich dimensionalen Vektorraum bez. einer gegebenen Basis dargestellt werden kann wie in (10.1) (analog, wie dies unten geschehen wird, Übung).

In diesem Kapitel wird es meist um Sesquilinearformen gehen. Dies sind ähnliche Objekte, die sich aber in einem wesentlichen Punkt von Bilinearformen unterscheiden.

**10.1 Definition** Sei  $V$  ein Vektorraum über  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ . Eine Abbildung  $Q : V \times V \rightarrow \mathbb{K}$  heißt eine Sesquilinearform auf  $V$ , falls  $Q$  antilinear im ersten Argument und linear im zweiten Argument ist:

$$Q(v + \lambda w, v' + \lambda' w') = Q(v, v') + \bar{\lambda} Q(w, v') + \lambda' Q(v, w') + \bar{\lambda} \lambda' Q(w, w') .$$

Die Menge aller Sesquilinearformen auf  $V$  wird mit  $\mathcal{SL}(V)$  bezeichnet. Des Weiteren wird definiert:

- (i)  $Q$  ist hermitisch genau dann, wenn  $Q(v, w) = \overline{Q(w, v)}$  für  $v, w \in V$ .
- (ii)  $Q$  ist anti-hermitisch genau dann, wenn  $Q(v, w) = -\overline{Q(w, v)}$  für  $v, w \in V$ .
- (iii)  $Q$  ist positiv definit genau dann, wenn  $Q(v, v) > 0$  für  $v \in V \setminus \{0\}$ .
- (iv)  $Q$  ist positiv semi-definit genau dann, wenn  $Q(v, v) \geq 0$  für  $v \in V$ .
- (v) Analog werden negative Definitheit und Semidefinitheit von  $Q$  definiert.
- (vi)  $Q$  ist nicht ausgeartet genau dann, wenn zu jedem  $v \in V \setminus \{0\}$  ein  $w \in V$  existiert mit  $Q(v, w) \neq 0$ . Anderenfalls heißt  $Q$  ausgeartet.
- (vii) Wenn  $\mathcal{B} = (b_1, \dots, b_N)$  eine geordnete Basis von  $V$  ist, so ist die darstellende Matrix von  $Q$

$$Q_{\mathcal{B}} = (Q(b_n, b_m))_{n,m=1,\dots,N} .$$

**10.2 Bemerkungen** 1. Falls  $\mathbb{K} = \mathbb{R}$ , fallen die Begriffe Bilinearform and Sequilinearform zusammen.

2. Skalarprodukte sind immer auch hermitische Sesquilinearformen. Hinzu kommen noch die Positivität und Nichtentartung (d.h. sie sind nicht ausgeartet). Über Positivität und den Zusammenhang zu positiver Definitheit wird weiter unten noch kommentiert.

3. Die Bezeichnung geht zurück auf *sesqui = semis que = anderthalb*.

4. Wenn versehen mit

$$(Q + \lambda Q')(v, w) = Q(v, w) + \lambda Q'(v, w) ,$$

ist  $\mathcal{SL}(V)$  ein  $\mathbb{K}$ -Vektorraum. ◇

**10.3 Beispiel** Sei  $V = \mathbb{K}^N$  und  $A \in \text{Mat}(N \times N, \mathbb{K})$ . Dann ist  $Q : V \times V \rightarrow \mathbb{K}$  definiert durch

$$Q(v, w) = \langle v | Aw \rangle = \sum_{n,m=1}^N \bar{v}_n A_{n,m} w_m , \quad v = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} , \quad w = \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix} ,$$

eine Sesquilinearform. ◇

Der folgende Satz besagt, dass jede Sesquilinearform auf einem endlichdimensionalen Vektorraum in Koordinaten durch dieses Beispiel gegeben ist.

**10.4 Satz** Sei  $\mathcal{B} = (b_1, \dots, b_N)$  eine geordnete Basis von  $V$  und  $I_{\mathcal{B}} : V \rightarrow \mathbb{K}^N$  die zugehörige Koordinatenabbildung. Dann ist

$$Q(v, w) = \langle I_{\mathcal{B}}(v) | Q_{\mathcal{B}} I_{\mathcal{B}}(w) \rangle ,$$

wobei auf der rechten Seite das Standardskalarprodukt auf  $\mathbb{K}^N$  verwandt wird. Die Abbildung  $Q \in \mathcal{SL}(V) \mapsto Q_{\mathcal{B}} \in \text{Mat}(N \times N, \mathbb{K})$  ist ein Vektorraumisomorphismus. Es gilt:

- (i)  $Q$  ist hermitisch genau dann, wenn  $Q_{\mathcal{B}} = (Q_{\mathcal{B}})^*$  selbstadjungiert (auch hermitisch genannt).
- (ii)  $Q$  ist anti-hermitisch genau dann, wenn  $Q_{\mathcal{B}} = -(Q_{\mathcal{B}})^*$  antiselbstadjungiert.
- (iii)  $Q$  ist positiv definit genau dann, wenn  $Q_{\mathcal{B}} > 0$ .
- (iv)  $Q$  ist positiv semi-definit genau dann, wenn  $Q_{\mathcal{B}} \geq 0$ .
- (v) Wenn  $\mathcal{C}$  eine weitere geordnete Basis von  $V$  ist und  $M = I_{\mathcal{B}}(I_{\mathcal{C}})^{-1}$  der zugehörige Basiswechsel, dann ist

$$Q_{\mathcal{C}} = M^* Q_{\mathcal{B}} M .$$

**Beweis.** Die Koordinatenabbildung ist gegeben durch

$$I_{\mathcal{B}}(v) = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} , \quad \text{für } v = \sum_{n=1}^N v_n b_n ,$$

und analog für  $I_{\mathcal{B}}(w)$ . Einsetzen ergibt also

$$\langle I_{\mathcal{B}}(v) | Q_{\mathcal{B}} I_{\mathcal{B}}(w) \rangle_{\mathcal{B}} = \sum_{n,m=1}^N \overline{v_n} (Q_{\mathcal{B}})_{n,m} w_m = \sum_{n,m=1}^N \overline{v_n} Q(b_n, b_m) w_m = Q(v, w) .$$

Die Vektorraumisomorphie ist offensichtlich, denn die Zuordnung ist linear und jede Matrix  $A \in \text{Mat}(N \times N, \mathbb{K})$  definiert eine Sesquilinearform  $Q(v, w) = \langle I_{\mathcal{B}}(v) | A I_{\mathcal{B}}(w) \rangle$ . Nun können alle weiteren Eigenschaften abgelesen werden. Zu (i):

$$\overline{Q(v, w)} = \langle Q_{\mathcal{B}} I_{\mathcal{B}}(w) | I_{\mathcal{B}}(v) \rangle = \langle I_{\mathcal{B}}(w) | (Q_{\mathcal{B}})^* I_{\mathcal{B}}(v) \rangle .$$

Dies ist nun für alle  $v, w$  gleich  $Q(w, v) = \langle I_{\mathcal{B}}(w) | Q_{\mathcal{B}} I_{\mathcal{B}}(v) \rangle$  genau dann, wenn  $Q_{\mathcal{B}} = (Q_{\mathcal{B}})^*$ . Analog folgt (ii). (iii) und (iv) sind Umformulierungen von Satz 9.8. Letztendlich zeigt

$$\langle I_{\mathcal{B}}(v) | Q_{\mathcal{B}} I_{\mathcal{B}}(w) \rangle = \langle I_{\mathcal{B}}(I_{\mathcal{C}})^{-1} I_{\mathcal{C}}(v) | Q_{\mathcal{B}} I_{\mathcal{B}}(I_{\mathcal{C}})^{-1} I_{\mathcal{C}}(w) \rangle = \langle I_{\mathcal{C}}(v) | M^* Q_{\mathcal{B}} M I_{\mathcal{C}}(w) \rangle$$

auch den letzten Punkt (v). □

**10.5 Definition** Zwei quadratische Matrizen  $A, B \in \text{Mat}(N \times N, \mathbb{K})$  heißen hermitisch kongruent genau dann, wenn eine invertierbare Matrix  $M \in \text{Mat}(N \times N, \mathbb{K})$  existiert, so dass

$$A = M^* B M .$$

Außerdem heißen  $A, B \in \text{Mat}(N \times N, \mathbb{K})$  kongruent genau dann, wenn  $A = M^t B M$  für eine invertierbare Matrix  $M \in \text{Mat}(N \times N, \mathbb{K})$ .

**10.6 Bemerkung** Kongruenz und hermitische Kongruenz sind Äquivalenzrelationen. ◇

**10.7 Satz** (Normalform für hermitische Sesquilinearformen) Sei  $Q$  eine hermitische Sesquilinearform auf einem endlich dimensionalen Vektorraum  $V$  über  $\mathbb{K}$  der Dimension  $N$ . Dann gibt es eine geordnete Basis  $\mathcal{B}$ , so dass  $Q_{\mathcal{B}}$  diagonal ist mit Diagonaleinträgen in  $\{1, 0, -1\}$ , d.h.

$$Q_{\mathcal{B}} = \begin{pmatrix} \mathbf{1}_{N_+} & 0 & 0 \\ 0 & 0_{N_0} & 0 \\ 0 & 0 & -\mathbf{1}_{N_-} \end{pmatrix},$$

wobei  $N_+ + N_0 + N_- = N$ .

**Beweis.** Zunächst sei  $\mathbb{K} = \mathbb{C}$ . Sei  $\mathcal{C} = (c_1, \dots, c_N)$  eine beliebige geordnete Basis von  $V$ . Dann ist nach Satz 10.4 die darstellende Matrix  $Q_{\mathcal{C}}$  selbstadjungiert. Nach dem Spektralsatz (Satz 7.33) gibt es eine unitäre Matrix  $U$  dergestalt, dass

$$U^* Q_{\mathcal{C}} U = \begin{pmatrix} z'_1 & & \\ & \ddots & \\ & & z'_N \end{pmatrix},$$

wobei die  $z'_n$  reell sind. Zudem kann  $U$  so gewählt werden, dass die Eigenwerte geordnet sind gemäß  $z'_n \geq z'_{n+1}$ . Nun seien  $N_+$  und  $N_-$  die Anzahl der positiven und negativen Eigenwerte, und  $N_0 = N - N_+ - N_-$ . Also  $z'_{N_++1} = \dots = z'_{N_++N_0} = 0$  und es gilt

$$\begin{pmatrix} z'_1 & & \\ & \ddots & \\ & & z'_N \end{pmatrix} = D \begin{pmatrix} \mathbf{1}_{N_+} & 0 & 0 \\ 0 & 0_{N_0} & 0 \\ 0 & 0 & -\mathbf{1}_{N_-} \end{pmatrix} D,$$

wobei die invertierbare, selbstadjungierte Diagonalmatrix  $D$  gegeben ist durch

$$D = \begin{pmatrix} |z'_1|^{\frac{1}{2}} & & & & & & \\ & \ddots & & & & & \\ & & |z'_{N_+}|^{\frac{1}{2}} & & & & \\ & & & \mathbf{1}_{N_0} & & & \\ & & & & |z'_{N_++N_0+1}|^{\frac{1}{2}} & & \\ & & & & & \ddots & \\ & & & & & & |z'_N|^{\frac{1}{2}} \end{pmatrix}.$$

Also gilt

$$(UD^{-1})^* Q_{\mathcal{C}} (UD^{-1}) = \begin{pmatrix} \mathbf{1}_{N_+} & 0 & 0 \\ 0 & 0_{N_0} & 0 \\ 0 & 0 & -\mathbf{1}_{N_-} \end{pmatrix}.$$

Nun verwenden wir  $M = UD^{-1}$  als Basiswechsel wie in Satz 10.4(v), d.h.  $\mathcal{B} = \mathcal{C}M$ . Dies beendet den Beweis im Fall  $\mathbb{K} = \mathbb{C}$ . Falls  $\mathbb{K} = \mathbb{R}$ , so wird die obige Basis  $\mathcal{C}$  in  $V$  gewählt (also reell). Dann ist  $Q_{\mathcal{C}}$  reell und die unitäre Matrix  $U$  kann gemäß Satz 7.35 auch reell gewählt werden. Da auch  $D$  reell ist, ist letztendlich auch  $M$  reell und somit  $\mathcal{B}$  eine Basis von  $V$  (und nicht seiner Komplexifizierung). □

Der erste Teil des obigen Beweises (beginne lediglich mit einer beliebigen Orthonormalbasis  $\mathcal{C}$ ) liefert folgende Aussage.



**10.8 Satz** (Hauptachsentransformation für hermitesche Sesquilinearformen) Sei  $V$  ein endlich dimensionaler Vektorraum  $V$  über  $\mathbb{K}$  versehen mit einem Skalarprodukt  $\langle \cdot | \cdot \rangle$ . Für jede hermitesche Sesquilinearform  $Q$  auf  $V$  existiert dann eine Orthonormalbasis  $\mathcal{B}$ , so dass  $Q_{\mathcal{B}}$  diagonal ist, d.h.

$$Q_{\mathcal{B}} = \begin{pmatrix} z'_1 & & \\ & \ddots & \\ & & z'_N \end{pmatrix}.$$

**10.9 Satz** (Trägheitssatz von Sylvester) Sei  $Q$  eine hermitesche Sesquilinearform auf einem endlich dimensionalen Vektorraum  $V$  über  $\mathbb{K}$  der Dimension  $N$ . Die Zahlen  $N_+$ ,  $N_0$  und  $N_-$ , die in Satz 10.7 auftauchen, sind auch gegeben durch

$$\begin{aligned} N_+ &= \max\{\dim(U) \mid U \text{ Unterraum mit } Q(u, u) > 0 \text{ für alle } u \in U \setminus \{0\}\}, \\ N_0 &= \max\{\dim(U) \mid U \text{ Unterraum mit } Q(v, u) = 0 \text{ für alle } v \in V \text{ und } u \in U\}, \\ N_- &= \max\{\dim(U) \mid U \text{ Unterraum mit } Q(u, u) < 0 \text{ für alle } u \in U \setminus \{0\}\}. \end{aligned}$$

Insbesondere sind die Anzahlen der positiven, verschwindenden und negativen Eigenwerte der Matrix  $M^*Q_{\mathcal{B}}M$  unabhängig von der Wahl der invertierbaren Matrix  $M$ .

**Beweis.** Wir betrachten zunächst  $N_+$  und bezeichnen das Maximum mit  $\max$ . Zuerst wird die Ungleichung  $N_+ \leq \max$  gezeigt. Nach dem Fischer Prinzip (Satz 9.5) gilt für den kleinsten positiven Eigenwert  $z'_{N_+}$  von  $Q_{\mathcal{B}}$

$$z'_{N_+} = \max_{\dim(\hat{U})=N_+} \min_{v \in \hat{U}, \|v\|=1} \langle v | Q_{\mathcal{B}}v \rangle.$$

Also existiert ein Unterraum  $\hat{U} \subset \mathbb{C}^N$  mit  $\dim(\hat{U}) = N_+$ , auf welchem  $\langle v | Q_{\mathcal{B}}v \rangle > 0$ . Also gilt für alle  $u \in U = (I_{\mathcal{B}})^{-1}(\hat{U})$ , dass  $Q(u, u) = \langle I_{\mathcal{B}}(u) | Q_{\mathcal{B}}I_{\mathcal{B}}(u) \rangle > 0$ . Somit  $N_+ \leq \max$ . Wenn nun aber  $N_+ < \max$ , so existiert ein Unterraum  $U$  mit  $\dim(U) = N_+ + 1$  mit  $Q(u, u) > 0$  für alle  $u \in U \setminus \{0\}$ . Somit wäre wieder nach dem Fischer Prinzip

$$z'_{N_++1} = \max_{\dim(\hat{U})=N_++1} \min_{v \in \hat{U}, \|v\|=1} \langle v | Q_{\mathcal{B}}v \rangle \geq \min_{v \in I_{\mathcal{B}}(U), \|v\|=1} \langle v | Q_{\mathcal{B}}v \rangle \geq \min_{u \in U, \|I_{\mathcal{B}}(u)\|=1} Q(u, u) > 0,$$

im Widerspruch zur Annahme. Analog wird bei  $N_-$  vorgegangen. Nun zu der Gleichheit für  $N_0$ . Die Bedingung  $Q(v, u) = \langle I_{\mathcal{B}}(v) | Q_{\mathcal{B}}I_{\mathcal{B}}(u) \rangle = 0$  für alle  $v \in V$  ist äquivalent zu  $Q_{\mathcal{B}}I_{\mathcal{B}}(u) = 0$ . Somit bleibt zu zeigen

$$N_0 = \max\{\dim(U) \mid U \text{ Unterraum mit } Q_{\mathcal{B}}I_{\mathcal{B}}(u) = 0 \text{ für alle } u \in U\},$$

d.h.  $N_0 = \dim(\text{Ker}(Q_{\mathcal{B}}))$ , was nach Satz 10.7 richtig ist. □

**10.10 Korollar** Sei  $Q$  eine hermitesche Sesquilinearform.

- (i)  $Q$  ist positiv definit genau dann, wenn  $N_0 = N_- = 0$ .
- (ii)  $Q$  ist positiv semi-definit genau dann, wenn  $N_- = 0$ .
- (iii)  $Q$  ist nicht ausgeartet genau dann, wenn  $N_0 = 0$ .

**10.11 Definition** Sei  $Q$  eine hermitesche Sesquilinearform.

- (i) Die Trägheit von  $Q$  ist dann das Tupel  $(N_+, N_0, N_-)$ .

(ii) Die Signatur von  $Q$  ist  $N_+ - N_-$ .

(iii) Der Witt Index von  $Q$  ist  $N_0 + \min\{N_+, N_-\}$ .

**10.12 Bemerkung** Die Trägheit (auf Englisch *inertia*) ist eine sogenannte Invariante der Sesquilinearform, ebenso Signatur und Witt Index. Oft wird auch die Trägheit schon Signatur genannt (was etwas verwirrend ist, aber alle Begriffe sind in der Tat nahe verwandt). Wenn  $N_0 = 0$ , dann wird oft auch  $(N_+, N_-)$  als Trägheit oder Signatur bezeichnet.  $\diamond$

**10.13 Beispiele** Sei  $V = \mathbb{R}^4$  die Minkovski Raum-Zeit, versehen mit dem Minkovski Produkt

$$Q(v, v') = xx' + yy' + zz' - tt', \quad v = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}, \quad v' = \begin{pmatrix} x' \\ y' \\ z' \\ t' \end{pmatrix}.$$

Dies ist eine Bilinearform und hier im Reellen auch eine Sesquilinearform. Wenn  $\mathcal{B}$  die Standardbasis ist, so

$$Q_{\mathcal{B}} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix}.$$

Also ist die Signatur  $(3, 1)$ . In diesem Zusammenhang wird  $Q$  auch als indefinite Metrik oder indefinites Skalarprodukt bezeichnet (weil sie weder positiv noch negativ definit ist).

Nun soll ein Beispiel für eine anti-hermitische Bilinearform angeführt werden. Sei  $V = \mathbb{R}^{2N}$ . Vektoren  $v \in V$  werden mit  $v = \begin{pmatrix} u \\ w \end{pmatrix}$  bezeichnet, wobei  $u, w \in \mathbb{R}^N$ . Wenn  $v' = \begin{pmatrix} u' \\ w' \end{pmatrix}$ , dann ist

$$Q(v, v') = \langle w|u' \rangle - \langle u|w' \rangle,$$

in der Tat eine anti-hermitische Bilinearform. Bzgl. der Standardbasis  $\mathcal{B}$  des  $\mathbb{R}$  ist dann

$$Q_{\mathcal{B}} = \begin{pmatrix} 0 & -\mathbf{1}_N \\ \mathbf{1}_N & 0 \end{pmatrix}.$$

Diese Bilinearform heißt auch die symplektische Form. Es ist möglich zu zeigen (Übung), dass anti-hermitische Bilinearformen auf dem  $\mathbb{R}^{2N}$  immer darstellende Matrizen von dieser Form haben.  $\diamond$

Wie oben schon angedeutet, werden hermitische Sesquilinearformen auch als sogenannte indefinite Skalarprodukte angesehen. Dies legt nahe, auch geometrische Begriffsbildungen wie Orthogonalität zu übertragen. Dies ist Teil der nächsten Definition.

**10.14 Definition** Sei  $Q$  eine hermitische oder anti-hermitische Sesquilinearform auf einem Vektorraum  $V$  über  $\mathbb{K}$ .

(i) Zwei Vektoren  $u, v \in V$  heißen  $Q$ -orthogonal genau dann, wenn  $Q(u, v) = 0$ .

(ii) Zwei Unterräume  $U$  und  $U'$  heißen  $Q$ -orthogonal genau dann, wenn  $Q(u, u') = 0$  für alle  $u \in U$  und  $u' \in U'$ . Kurzschreibweise hierfür ist  $U \perp_Q U'$ .

(iii) Das  $Q$ -orthogonale Komplement eines Unterraumes  $U$  ist

$$U^{\perp_Q} = \{v \in V \mid Q(v, u) = 0 \text{ für alle } u \in U\}.$$

- (iv) Ein Unterraum  $U$  ist  $Q$ -isotrop genau dann, wenn  $Q(u, u) = 0$  für alle  $u \in U$ .
- (v) Ein Unterraum  $U$  ist  $Q$ -Lagrange genau dann, wenn  $U$  maximal als  $Q$ -isotroper Unterraum ist, d.h. nicht Teilmenge eines echt größeren  $Q$ -isotropen Unterraumes ist.
- (vi) Ein Unterraum  $U$  heißt  $Q$ -positiv genau dann, wenn  $Q(u, u) > 0$  für alle  $u \in U \setminus \{0\}$ . Analog sind  $Q$ -negative Unterräume definiert.

**10.15 Bemerkung** Das Wort *isotrop* ist eine Erfindung der Mathematiker des 19ten Jahrhundert, basierend auf dem Griechischen für *gleich Drehung*. Der Ursprung und die Bedeutung des Wortes ist im Buch von Brieskorn [Bri] detailliert diskutiert. Dort wird obige  $Q$ -Isotropie allerdings mit Totalisotropie bezeichnet (was recht unüblich ist).  $\diamond$

**10.16 Satz** Sei  $Q$  eine hermitische Sesquilinearform auf einem Vektorraum  $V$  über  $\mathbb{K}$ , und sei  $U$  ein Unterraum. Dann:

- (i)  $U \subset (U^{\perp Q})^{\perp Q}$
- (ii)  $U$  ist  $Q$ -isotrop genau dann, wenn  $U \subset U^{\perp Q}$ .
- (iii) Wenn  $U$   $Q$ -Lagrange ist, so ist  $\dim(U)$  gleich dem Witt-Index  $N_0 + \min\{N_+, N_-\}$ .

Wenn  $Q$  zudem nicht entartet ist, so gilt:

- (iv)  $\dim(U) + \dim(U^{\perp Q}) = \dim(V)$
- (v)  $U = (U^{\perp Q})^{\perp Q}$

**Beweis.** (i) Wenn  $u \in U$ , dann gilt nach Definition  $Q(u, v) = 0$  für alle  $v \in U^{\perp Q}$ . Also, wiederum nach Dimension  $u \in (U^{\perp Q})^{\perp Q}$ . (ii) Dies ist offensichtlich nach der Definition. (iii) Sei  $U_+$  ein Unterraum der Dimension  $N_+$ , auf dem  $Q$  positiv ist (vgl. Satz 10.9). Wenn  $U$  isotrop ist, so kann es keinen Vektor aus  $U_+$  enthalten. Also ist  $U$  nach dem Dimensionssatz von Dimension kleiner oder gleich  $N - N_+$  (denn sonst gäbe es einen nicht-trivialen Schnitt). Analog,  $\dim(U) \leq N - N_-$ . Zusammen

$$\dim(U) \leq \min\{N - N_+, N - N_-\} = \min\{N_- + N_0, N_+ + N_0\} = N_0 + \min\{N_+, N_-\}.$$

Nun werden wir zeigen, dass es einen isotropen Unterraum gibt, der diese maximale Dimension annimmt. Dies wird unter Verwendung der Normalform aus Satz 10.7 geschehen, also in der Darstellung bez. einer Basis. Der Einfachheit nehmen wir an, dass  $N_+ \leq N_-$ . Dann wählen wir

$$u_n = \begin{cases} e_n + e_{n+N_+ + N_0}, & n = 1, \dots, N_+, \\ e_n, & n = N_+ + 1, \dots, N_+ + N_0. \end{cases}$$

Tatsächlich spannen diese Vektoren  $u_1, \dots, u_{N_+ + N_0}$  einen Unterraum  $U \subset \mathbb{C}^N$  auf, der isotrop und der gewünschten Dimension  $N_+ + N_0$  ist. Dann hat auch  $(I_B)^{-1}(U) \subset V$  diese Eigenschaften. Außerdem kann gezeigt werden, dass jeder Lagrange'sche Unterraum (maximal isotroper Unterraum) von dieser Dimension ist, weil er sonst nämlich vergrößert werden kann.

(iv) Wiederum braucht dies lediglich für die Normalform nachgewiesen werden, denn die Koordinatenabbildung  $I_B$  ist ja ein Isomorphismus. Nun gilt aber

$$\begin{aligned} U^{\perp Q} &= \{v \in V \mid \langle u, Q_B v \rangle = 0 \quad \forall u \in U\} \\ &= \{v \in V \mid \langle Q_B u, v \rangle = 0 \quad \forall u \in U\} = (Q_B U)^{\perp} = (Q_B)^{-1} U^{\perp}, \end{aligned}$$

wobei im letzten Schritt die Invertierbarkeit und Selbstadjungiertheit von  $Q_B$  verwandt wurde, sowie die Identität  $(TU)^\perp = (T^*)^{-1}U^\perp$ , die auf beliebigen Vektorräumen mit Skalarprodukt und invertierbares  $T$  gilt (dies sei als Übung verifiziert). Wegen der vorausgesetzten Invertierbarkeit von  $Q_B$  folgt also

$$\dim(\mathbb{K}^N) = \dim(U) + \dim(U^\perp) = \dim(U) + \dim((Q_B)^{-1}U^\perp) = \dim(U) + \dim(U^{\perp Q}),$$

wobei die erste Gleichheit aus Satz 7.40(ii) folgt. Für (v) kann nun wie folgt argumentiert werden:

$$(U^{\perp Q})^{\perp Q} = (Q_B(Q_B)^{-1}U^\perp)^\perp = (U^\perp)^\perp = U,$$

wobei im letzten Schritt Satz 7.40(iii) verwandt wurde. □

**10.17 Bemerkung** Analoge Aussagen gelten auch für anti-hermitische Bilinearformen, also insbesondere die symplektische Form. ◇

**10.18 Definition** Sei  $Q$  eine hermitische oder anti-hermitische Sesquilinearform auf einem Vektorraum  $V$  über  $\mathbb{K}$ . Eine lineare Abbildung  $T \in \mathcal{L}(V)$  heißt  $Q$ -unitär genau dann, wenn

$$Q(Tv, Tw) = Q(v, w), \quad v, w \in V.$$

**10.19 Satz** Die  $Q$ -unitären linearen Abbildungen bilden eine Gruppe bez. der Hintereinanderausführung.

**Beweis.** Seien  $S, T \in \mathcal{L}(V)$  beide  $Q$ -unitär. Dann gilt

$$Q(STv, STw) = Q(Tv, Tw) = Q(v, w),$$

für alle  $v, w \in V$ . □

**10.20 Beispiel** Sei  $V = \mathbb{C}^{N+M}$  und

$$Q(v, w) = \langle v | \begin{pmatrix} \mathbf{1}_N & 0 \\ 0 & -\mathbf{1}_M \end{pmatrix} | w \rangle.$$

Diese hermitische Sesquilinearform ist schon in ihrer Normalform gegeben und heißt die Lorentzform der Signatur  $(N, M)$ . Die zugehörige Gruppe der  $Q$ -unitären Matrizen wird mit  $U(N, M)$  bezeichnet und heißt verallgemeinerte Lorentzgruppe der Signatur  $(N, M)$ . Wenn man

$$J = \begin{pmatrix} \mathbf{1}_N & 0 \\ 0 & -\mathbf{1}_M \end{pmatrix}$$

setzt, dann ist für eine Matrix  $A$  der Größe  $N + M$

$$Q(Av, Aw) = \langle Av | JAw \rangle = \langle v | (A^* J A) w \rangle, \quad Q(v, w) = \langle v | J w \rangle.$$

Also gilt

$$A \in U(N, M) \iff A^* J A = J. \tag{10.2}$$

Wenn  $A$  reell ist und die Signatur  $(N, M) = (3, 1)$ , heißt  $A$  auch eine Lorentztransformation. ◇

**10.21 Beispiel** Sei  $V = \mathbb{R}^{2N}$  und

$$Q(v, w) = \langle v | \begin{pmatrix} 0 & -\mathbf{1}_N \\ \mathbf{1}_N & 0 \end{pmatrix} | w \rangle .$$

Diese anti-hermitesche Sesquilinearform heißt die symplektische Form (siehe oben). Die zugehörige Gruppe der  $Q$ -unitären Matrizen wird mit  $\text{SP}(2N, \mathbb{R})$  bezeichnet und heißt die symplektische Gruppe der Dimension  $2N$ . Eine Matrix  $A \in \text{Mat}(2N \times 2N, \mathbb{R})$  ist symplektisch genau dann, wenn

$$A^t J A = J, \quad J = \begin{pmatrix} 0 & -\mathbf{1}_N \\ \mathbf{1}_N & 0 \end{pmatrix} .$$

Dies sieht analog wie (10.2) aus und kann auch genauso nachgewiesen werden. Der wesentliche Unterschied ist natürlich die Form der Matrix  $J$ .  $\diamond$

**10.22 Bemerkung** Die verallgemeinerte Lorentzgruppe und die symplektische Gruppe sind 2 von Hermann Weyl's 10 sogenannten Klassischen Gruppen. Zwei weitere sind die allgemeinen linearen Gruppen  $\text{GL}(N, \mathbb{C})$  und  $\text{GL}(N, \mathbb{R})$ . Alle weiteren Klassischen Gruppen sind auch Untergruppen der allgemeinen linearen Gruppen, die Symmetrien ähnlich wie (10.2) erfüllen.  $\diamond$

**10.23 Satz** Sei  $A \in \text{U}(N, M)$  und  $J$  wie in (10.2). Dann gilt:

(i) Das Spektrum hat eine Spiegelsymmetrie am Einheitskreis:

$$\text{Spec}(A) = \overline{(\text{Spec}(A))}^{-1} .$$

(ii) Wenn  $Av = zv$  und  $Aw = \zeta w$  mit  $\bar{z}\zeta \neq 1$ , dann  $v \perp_J w$ .

(iii) Die gesamte algebraische Multiplizität aller Eigenwerte auf dem Einheitskreis ist mindestens  $|N - M|$  (hier taucht also die Signatur auf).

**Beweis.** (i) Aus (10.2) und  $J^{-1} = J$  folgt  $A^{-1} = J A^* J$ . Also gilt für das charakteristische Polynom

$$p_A(z) = \det(z \mathbf{1} - A) = \overline{\det(J) \det(\bar{z} \mathbf{1} - A^*) \det(J)} = \overline{\det(\bar{z} \mathbf{1} - A^{-1})} = \overline{p_{A^{-1}}(\bar{z})} .$$

Somit

$$\text{Spec}(A) = \overline{\text{Spec}(A^{-1})} = \overline{\text{Spec}(A)}^{-1} ,$$

wobei im letzten Schritt der spektrale Abbildungssatz (Satz 8.15) verwandt wurde.

(ii) Die Aussage  $v \perp_J w$  ist äquivalent zu  $\langle v | J w \rangle = 0$ . Nun gilt

$$\langle v | J w \rangle = \frac{1}{\bar{z}\zeta} \langle Av | J Aw \rangle = \frac{1}{\bar{z}\zeta} \langle v | A^* J Aw \rangle = \frac{1}{\bar{z}\zeta} \langle v | J w \rangle ,$$

so dass aus der Annahme  $\bar{z}\zeta \neq 1$  tatsächlich  $\langle v | J w \rangle = 0$  folgt.

(iii) Wir beweisen dies lediglich unter der vereinfachenden Annahme, dass keine Jordanblöcke vorliegen (ansonsten muss (ii) auch für verallgemeinerte Eigenvektoren nachgewiesen werden). Sei  $U$  der Unterraum, der von allen Eigenvektoren zu Eigenwerten außerhalb des Einheitskreises (Betrag größer als 1) aufgespannt wird. Nach (ii) und der Zusatzannahme ist dies ein isotroper Unterraum. Seine Dimension also nach Satz 10.16 beschränkt durch den Witt-Index  $\min\{N, M\}$ . Genauso ist die Multiplizität aller Eigenwerte vom Betrag kleiner als 1 beschränkt durch  $\min\{N, M\}$ . Also liegen  $N + M - 2 \min\{N, M\} = |N - M|$  Eigenwerte (gezählt mit ihrer Multiplizität) auf dem Einheitskreis.  $\square$

Nun sollen die obigen Konzepte verwandt werden, um Quadriken zu studieren.

**10.24 Definition** Ein reelles quadratisches Polynom in  $N$  Variablen  $x_1, \dots, x_N$  ist ein Polynom  $p$  von der Gestalt

$$P(x_1, \dots, x_N) = \sum_{n,m=1}^N p_{n,m} x_n x_m + \sum_{n=1}^N p_n x_n + p, \quad (10.3)$$

wobei  $p_{n,m}, p_n, p \in \mathbb{R}$ . Eine reelle Quadrik  $Q$  im  $\mathbb{R}^N$  ist die Nullstellenmenge eines reellen quadratischen Polynomes  $P$  in  $N$  Variablen:

$$Q = \{(x_1, \dots, x_N) \in \mathbb{R}^N \mid P(x_1, \dots, x_N) = 0\}.$$

**10.25 Satz** Zu jedem quadratischen Polynom  $P$  in  $N$  Variablen  $x_1, \dots, x_N$  gibt es eine selbstadjungierte reelle Matrix  $A = (a_{n,m})_{n,m=1,\dots,N}$ , einen Vektor  $a_0 \in \mathbb{R}^N$  und eine Zahl  $a_{0,0} \in \mathbb{R}$ , so dass

$$P(x_1, \dots, x_N) = \langle x | Ax \rangle + 2 \langle a_0 | x \rangle + a_{0,0}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}.$$

Es ist möglich dies umzuschreiben mit Hilfe von  $A' = (a_{n,m})_{n,m=0,\dots,N}$ :

$$P(x_1, \dots, x_N) = \langle x' | A' x' \rangle, \quad A' = \begin{pmatrix} a_{0,0} & (a_0)^t \\ a_0 & A \end{pmatrix}, \quad x' = \begin{pmatrix} 1 \\ x \end{pmatrix}.$$

**Beweis.** Ausgehend von der Darstellung (10.3) setze  $a_{n,m} = \frac{1}{2}(p_{n,m} + p_{m,n})$  und  $a_{0,n} = \frac{1}{2}p_n$  und  $a_{0,0} = p$ . Da  $x_n x_m = x_m x_n$ , gilt dann die erste Darstellungsformel. Die zweite folgt nach Ausmultiplizieren, weil  $\langle a_0 | x \rangle = \langle x | a_0 \rangle$ .  $\square$

Als Erstes wird das Transformationsverhalten von Quadriken unter sogenannten affinen Abbildungen untersucht.

**10.26 Definition** Sei  $V$  ein Vektorraum über  $\mathbb{K}$ . Eine Abbildung  $F : V \rightarrow V$  heißt affin genau dann, wenn es eine lineare Abbildung  $T \in \mathcal{L}(V)$  und einen Vektor  $b$  gibt, so dass

$$F(v) = Tv + b.$$

- (i) Falls  $T = \mathbf{1}_V$ , heißt  $F$  eine Translation.
- (ii) Eine Affinität ist eine bijektive affine Abbildung.
- (iii) Falls  $V$  mit einem Skalarprodukt versehen ist und  $T$  orthogonal ist, so heißt  $F$  eine euklidische Bewegung.

Zugehörig zu einer affinen Abbildung wird eine lineare Abbildung  $F' \in \mathcal{L}(\mathbb{K} \oplus V)$  definiert via

$$F' = \begin{pmatrix} 1 & 0 \\ b & T \end{pmatrix}.$$

Hierbei ist  $\mathbb{K} \oplus V$  der  $\mathbb{K}$ -Vektorraum (genannt die direkte Summe), gegeben durch die Menge  $\mathbb{K} \times V$  versehen mit

$$\begin{pmatrix} a \\ v \end{pmatrix} + \lambda \begin{pmatrix} a' \\ v' \end{pmatrix} = \begin{pmatrix} a + \lambda a' \\ v + \lambda v' \end{pmatrix}, \quad \lambda \in \mathbb{K}.$$

**10.27 Bemerkung** Das Inverse einer Affinität  $F(v) = Tv + b$  ist gegeben durch  $F^{-1}(v) = T^{-1}v - T^{-1}b$ , also wieder eine Affinität. Also bilden die Affinitäten eine Gruppe. Die Translationen bilden eine Untergruppe, ebenso wie die invertierbaren linearen Abbildungen auf  $V$ .  $\diamond$

Der folgende Satz besagt im Wesentlichen, dass affine Abbildungen als lineare Abbildungen auf dem vergrößerten Vektorraum  $\mathbb{K} \oplus V$  aufgefasst werden können.

**10.28 Satz** Seien  $F$  und  $G$  affine Abbildungen auf  $V$  und zu  $v \in V$  setze  $v' = \begin{pmatrix} 1 \\ v \end{pmatrix} \in \mathbb{K} \oplus V$ . Dann gelten die folgenden Identitäten:

- (i)  $F'v' = F(v)'$ , wobei  $F'v'$  die übliche Matrix-Vektormultiplikation in  $\mathbb{K} \oplus V$  ist.
- (ii)  $(F \circ G)' = F'G'$
- (iii)  $(F^{-1})' = (F')^{-1}$ , falls  $F$  eine Affinität ist.

**Beweis.** In der Tat,

$$F'v' = \begin{pmatrix} 1 & 0 \\ b & T \end{pmatrix} \begin{pmatrix} 1 \\ v \end{pmatrix} = \begin{pmatrix} 1 \\ b + Tv \end{pmatrix} = F(v)' ,$$

und ähnlich

$$F'G' = \begin{pmatrix} 1 & 0 \\ b & T \end{pmatrix} \begin{pmatrix} 1 & 0 \\ c & S \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ b + Tc & TS \end{pmatrix} = (F \circ G)' .$$

Hieraus folgt auch (iii).  $\square$

**10.29 Bemerkung** Die Gruppe der Affinitäten wird über die Zuordnung  $F \mapsto F'$  mit der Untergruppe der invertierbaren linearen Abbildung auf  $\mathbb{K} \oplus V$  von der oberen Dreiecksform

$$\begin{pmatrix} 1 & 0 \\ b & T \end{pmatrix}$$

identifiziert. Die euklidischen Bewegungen bilden eine Untergruppe der Affinitäten.  $\diamond$

**10.30 Satz** Sei  $\mathcal{Q}$  eine Quadrik im  $\mathbb{R}^N$  und  $F$  eine Affinität auf dem  $\mathbb{R}^N$ . Dann ist auch das Bild  $F(\mathcal{Q})$  eine Quadrik.

**Beweis.** Nach Satz 10.25 ist die Quadrik von der Gestalt

$$\mathcal{Q} = \{x \in \mathbb{R}^N \mid \langle x' | A' x' \rangle = 0\} ,$$

wobei  $A' = (A')^t$  eine quadratische Matrix der Größe  $N + 1$  ist. Somit unter Verwendung von Satz 10.28

$$\begin{aligned} F(\mathcal{Q}) &= \{F(x) \in \mathbb{R}^N \mid \langle x' | A' x' \rangle = 0\} \\ &= \{y \in \mathbb{R}^N \mid \langle (F^{-1}y)' | A' (F^{-1}y)' \rangle = 0\} \\ &= \{y \in \mathbb{R}^N \mid \langle (F^{-1})' y' | A' (F^{-1})' y' \rangle = 0\} \\ &= \{y \in \mathbb{R}^N \mid \langle y' | ((F')^{-1})^t A' (F')^{-1} y' \rangle = 0\} . \end{aligned}$$

Da  $((F')^{-1})^t A' (F')^{-1}$  wieder eine symmetrische Matrix ist, liegt also in der Tat wieder ein Quadrik vor.  $\square$

Das Ziel des folgenden Satzes ist es zu einer gegebenen Quadrik eine Affinität zu bestimmen, die sie in eine besonders einfache Form bringt.

**10.31 Satz** (Affine Normalformen von reellen Quadriken) Sei  $\mathcal{Q}$  eine Quadrik im  $\mathbb{R}^N$  von der Form

$$\mathcal{Q} = \{x \in \mathbb{R}^N \mid \langle x|Ax \rangle + 2\langle a_0|x \rangle + a_{0,0} = 0\} = \{x \in \mathbb{R}^N \mid \langle x'|A'x' \rangle = 0\}.$$

Die Signatur von  $A$  sei  $(k, N - k - K, K)$  und die Ränge  $r = \text{rk}(A)$  und  $r' = \text{rk}(A')$ , d.h.  $k + K = r$ . Dann existiert eine Affinität  $F$  auf dem  $\mathbb{R}^N$ , so dass  $F(\mathcal{Q})$  in einer der folgenden Normalformen ist:

$$\begin{aligned} \text{(I)} \quad & y_1^2 + \dots + y_k^2 - y_{k+1}^2 - \dots - y_r^2 = 0 && \text{falls } r = r', \\ \text{(II)} \quad & y_1^2 + \dots + y_k^2 - y_{k+1}^2 - \dots - y_r^2 = \pm 1 && \text{falls } r + 1 = r', \\ \text{(III)} \quad & y_1^2 + \dots + y_k^2 - y_{k+1}^2 - \dots - y_r^2 + 2y_{r+1} = 0 && \text{falls } r + 2 = r'. \end{aligned}$$

**Beweis.** Wir werden  $F$  als  $F'$  durch eine Hintereinanderausführung von mehreren Affinitäten  $F'_k$  explizit konstruieren. Zunächst sei  $M$  so wie in Satz 10.7 gewählt (bis auf das Vertauschen der beiden Blöcke  $\mathbf{1}_{N_-}$  und  $0_{N_0}$ ). Setze

$$F'_1 = \begin{pmatrix} 1 & 0 \\ 0 & M \end{pmatrix}.$$

Dann ist

$$(F'_1)^t A' F'_1 = \begin{pmatrix} a_{0,0} & c_+^t & c_-^t & c_0^t \\ c_+ & \mathbf{1}_k & 0 & 0 \\ c_- & 0 & -\mathbf{1}_K & 0 \\ c_0 & 0 & 0 & 0 \end{pmatrix},$$

wobei  $c_+ \in \mathbb{R}^k$ ,  $c_- \in \mathbb{R}^K$  und  $c_0 \in \mathbb{R}^{N-r}$ . Dann sei

$$F'_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -c_+ & \mathbf{1}_k & 0 & 0 \\ c_- & 0 & \mathbf{1}_K & 0 \\ 0 & 0 & 0 & \mathbf{1}_{N-r} \end{pmatrix}.$$

Dann gilt nach kurzer Rechnung

$$(F'_1 F'_2)^t A' (F'_1 F'_2) = \begin{pmatrix} \alpha & 0 & 0 & c_0^t \\ 0 & \mathbf{1}_k & 0 & 0 \\ 0 & 0 & -\mathbf{1}_K & 0 \\ c_0 & 0 & 0 & 0 \end{pmatrix},$$

wobei  $\alpha = a_{0,0} - c_+^t c_+ + c_-^t c_-$ .

Wenn nun  $\alpha = 0$  und  $c_0 = 0$ , so ist  $\text{rk}(A') = k + K = \text{rk}(A)$  und wir sind im Fall (I). Der Beweis ist dann schon beendet, denn mit  $F' = (F'_1 F'_2)^{-1}$  und  $y = (F')(x)$  gilt schon

$$\langle x'|A'x' \rangle = \langle y'|((F')^{-1})^t A' (F')^{-1} y' \rangle = \langle y'| \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \mathbf{1}_k & 0 & 0 \\ 0 & 0 & -\mathbf{1}_K & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} y' \rangle,$$

was die gewünschte Darstellung ist.

Wenn nun  $\alpha \neq 0$  und  $c_0 = 0$ , so ist  $\text{rk}(A') = k + K + 1 = \text{rk}(A) + 1$  und wir sind im Fall (II). Dann gilt mit

$$F'_3 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{|\alpha|} \mathbf{1}_N \end{pmatrix},$$



dass

$$(F'_1 F'_2 F'_3)^t A' (F'_1 F'_2 F'_3) = |\alpha| \begin{pmatrix} \operatorname{sgn}(\alpha) & 0 & 0 & 0 \\ 0 & \mathbf{1}_k & 0 & 0 \\ 0 & 0 & -\mathbf{1}_K & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

was bis auf den irrelevanten globalen Faktor  $|\alpha|$  (da wir die Nullstellenmenge betrachten) wieder die angegebene Normalform produziert.

Wenn nun  $c_0 \neq 0$ , so ist  $\operatorname{rk}(A') = k + K + 2 = \operatorname{rk}(A) + 2$  und wir sind im Fall (III). Sei  $T \in \operatorname{Gl}(N-r, \mathbb{R})$ , so dass  $T^t c_0 = e_1$  gilt (ansonsten ist  $T$  beliebig). Mit

$$F'_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \mathbf{1}_k & 0 & 0 \\ 0 & 0 & \mathbf{1}_K & 0 \\ 0 & 0 & 0 & T \end{pmatrix},$$

gilt dann

$$(F'_1 F'_2 F'_4)^t A' (F'_1 F'_2 F'_4) = \begin{pmatrix} \alpha & 0 & 0 & e_1^t \\ 0 & \mathbf{1}_k & 0 & 0 \\ 0 & 0 & -\mathbf{1}_K & 0 \\ e_1 & 0 & 0 & 0 \end{pmatrix}.$$

Im letzten Schritt setze dann

$$F'_5 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \mathbf{1}_k & 0 & 0 \\ 0 & 0 & \mathbf{1}_K & 0 \\ -\frac{\alpha}{2} e_1 & 0 & 0 & \mathbf{1}_{N-r} \end{pmatrix}.$$

Es folgt nach einer Rechnung in dem relevanten  $2 \times 2$  Block:

$$(F'_1 F'_2 F'_4 F'_5)^t A' (F'_1 F'_2 F'_4 F'_5) = \begin{pmatrix} 0 & 0 & 0 & e_1^t \\ 0 & \mathbf{1}_k & 0 & 0 \\ 0 & 0 & -\mathbf{1}_K & 0 \\ e_1 & 0 & 0 & 0 \end{pmatrix}.$$

Wiederum ergibt dies die angegebene Normalform. □

Nun können in Dimension  $N = 2$  alle möglichen Normalformen von Quadriken aufgelistet werden.

Typ	$r$	$r'$	$k$	Gleichung	Quadrik
I	0	0	0	$0 = 0$	Ebene $\mathbb{R}^2$
I	1	1	1	$x_1^2 = 0$	Gerade
I	2	2	1	$x_1^2 - x_2^2 = 0$	Geradenpaar mit Schnitt
I	2	2	2	$x_1^2 + x_2^2 = 0$	Punkt
II	1	2	1	$x_1^2 = 1$	paralleles Geradenpaar
II	2	3	1	$x_1^2 - x_2^2 = 1$	Hyperbel
II	2	3	2	$x_1^2 + x_2^2 = 1$	Kreis
III	1	3	1	$x_1^2 + 2x_2 = 0$	Parabel

Hierbei ist zu beachten, dass eine Ellipse im Satz 10.31 affin auf einen Kreis abgebildet wird, der also die Normalform der Ellipse ist. Die affinen Abbildungen erlauben es, die Hauptachsen zu normieren. Anders

ist die Situation, wenn nur euklidische Bewegungen (Komposition von Translationen und orthogonalen Abbildungen) verwandt werden. Dann können die Hauptachsen nicht normiert werden, und es gibt z.B. anstelle des Kreises eine Zweiparameterfamilie von Ellipsen (Thema einer Übung).

In Dimension  $N = 3$  ist die Liste der Normalformen länger, aber noch überschaubar.

Typ	$r$	$r'$	$k$	$K$	Gleichung	Quadrik
I	0	0	0	0	$0 = 0$	$\mathbb{R}^3$
I	1	1	1	0	$x_1^2 = 0$	Ebene
I	2	2	1	1	$x_1^2 - x_2^2 = 0$	Ebenenpaar mit Schnitt
I	2	2	2	0	$x_1^2 + x_2^2 = 0$	Gerade
I	3	3	2	1	$x_1^2 + x_2^2 - x_3^2 = 0$	Kreiskegel
I	3	3	3	0	$x_1^2 + x_2^2 + x_3^2 = 0$	Punkt
II	1	2	1	0	$x_1^2 = 1$	paralleles Ebenenpaar
II	2	3	1	0	$x_1^2 - x_2^2 = 1$	hyperbolischer Zylinder
II	2	3	2	0	$x_1^2 + x_2^2 = 1$	Kreiszyylinder
II	3	4	1	2	$x_1^2 - x_2^2 - x_3^2 = 1$	zweischaliges Hyperboloid
II	3	4	2	1	$x_1^2 + x_2^2 - x_3^2 = 1$	einschaliges Hyperboloid
II	3	4	3	0	$x_1^2 + x_2^2 + x_3^2 = 1$	Kugel
III	1	3	1	0	$x_1^2 + 2x_2 = 0$	parabolischer Zylinder
III	2	4	1	1	$x_1^2 - x_2^2 + 2x_3 = 0$	hyperbolisches Paraboloid
III	2	4	2	0	$x_1^2 + x_2^2 + 2x_3 = 0$	elliptisches Paraboloid

Zuletzt sei noch einmal darauf verwiesen, dass es auch Euklidische Normalformen von Quadriken gibt. Hier dürfen lediglich euklidische Bewegungen (Translationen und Rotationen) verwandt werden, um die Quadriken auf einen Normalform zu bringen. Es verbleiben dann noch die unterschiedlichen Längen der Hauptachsen. Z.B. ist die euklidische Normalform eines zweischaligen Hyperboloids

$$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} - \frac{x_3^2}{a_3^2} = b^2,$$

wobei  $a_1, a_2, a_3, b \in \mathbb{R}$  die Hauptachsenlängen sind. Dies erlaubt dann z.B. auch eine Kugel von einem Ellipsoid zu unterscheiden, welche die gleiche affine Normalform haben.

## 11 Matrixnormen

**11.1 Definition** Sei  $V$  ein Vektorraum über  $\mathbb{K}$ . Eine Abbildung  $\| \cdot \| : V \rightarrow [0, \infty)$  heißt Norm, genau dann wenn

- (i)  $\|\lambda v\| = |\lambda| \|v\|$  (Homogenität)
- (ii)  $\|v + w\| \leq \|v\| + \|w\|$  (Dreiecksungleichung)
- (iii)  $\|v\| = 0 \implies v = 0$  (Nicht-Entartung)

Dann ist  $(V, \|\cdot\|)$  ein normierter Vektorraum.

**11.2 Bemerkung** Auf einem normierten Vektorraum kann eine Metrik definiert werden durch

$$d(v, w) = \|v - w\| ,$$

welche  $V$  zu einem metrischen Raum macht, vgl. Analysis. Es ist dann möglich den Konvergenzbegriff bez. dieser Metrik einzuführen. Eine Folge  $(v_k)_{k \geq 1}$  von Vektoren  $v_k \in V$  konvergiert in  $(V, \|\cdot\|)$  gegen  $v \in V$  genau dann, wenn  $\lim_{k \rightarrow \infty} \|v_k - v\| = 0$ . Des Weiteren ist der Begriff der Cauchy-Folge verfügbar. Ein normierter Vektorraum  $V$ , in dem jede Cauchy-Folge konvergiert (d.h.  $V$  ist vollständig als metrischer Raum), heißt ein Banachraum.  $\diamond$

**11.3 Bemerkung** Ein Beispiel einer Norm haben wir schon kennengelernt. Wenn auf einem Vektorraum  $V$  ein Skalarprodukt  $\langle \cdot | \cdot \rangle$  gegeben ist, dann zeigt Satz 7.13, dass  $\|v\| = \sqrt{\langle v | v \rangle}$  eine Norm ist. Diese Norm heißt die euklidische Norm. Es gibt jedoch auch andere, wie folgende Beispiele zeigen.  $\diamond$

**11.4 Beispiel** Sei  $V = \mathbb{K}^N$ . Dann sind für  $p \geq 1$

$$\|v\|_p = \left( \sum_{n=1}^N |v_n|^p \right)^{\frac{1}{p}} , \quad \|v\|_\infty = \max_{n=1, \dots, N} |v_n| , \quad v = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} ,$$

Normen auf dem  $\mathbb{K}^N$ . Hierbei ist der Fall  $p = 2$  die schon bekannte euklidische Norm zum Standardskalarprodukt, d.h.  $\|v\|_2 = \langle v | v \rangle^{\frac{1}{2}}$ . Der Nachweis der Homogenität und Nicht-Entartung ist einfach, aber für den Beweis der Dreiecksungleichung wurde im Fall  $p = 2$  die Cauchy-Schwarz Ungleichung verwandt. Für  $p = \infty$  gilt

$$\|v+w\|_\infty = \max_{n=1, \dots, N} |v_n + w_n| \leq \max_{n=1, \dots, N} (|v_n| + |w_n|) \leq \max_{n=1, \dots, N} |v_n| + \max_{n=1, \dots, N} |w_n| = \|v\|_\infty + \|w\|_\infty .$$

Ähnlich kann im Fall  $p = 1$  vorgegangen werden. Die Dreiecksungleichung im Fall  $p \neq 2$  heißt die sogenannte Minkovski-Ungleichung, welche in Appendix B bewiesen wird. Des Weiteren sei auf einen Zusammenhang zwischen den Normen  $\|\cdot\|_\infty$  und  $\|\cdot\|$  hingewiesen:

$$\|v\|_\infty \leq \|v\|_2 \leq \sqrt{N} \|v\|_\infty . \tag{11.1}$$

Dies ist ein Beispiel von zwei Normen, die zueinander äquivalent genannt werden, weil nämlich die Konvergenzbegriffe zusammenfallen (eine Folge konvergiert bez. der einen Norm genau dann, wenn sie bez. der äquivalenten Norm konvergiert). In endlich dimensionalen Vektorräumen sind nun alle Normen äquivalent, wie folgender Satz zeigt.  $\diamond$

**11.5 Satz** Sei  $V$  ein endlich dimensionaler Vektorraum versehen mit zwei Normen  $\|\cdot\|$  und  $\|\cdot\|_\sim$ . Dann sind diese Normen äquivalent, d.h. es gibt eine Konstante  $C$  mit

$$\frac{1}{C} \|v\| \leq \|v\|_\sim \leq C \|v\| , \quad v \in V .$$

Als Vorbereitung für den Beweis zeigen wir zunächst ein Lemma.

**11.6 Lemma** Sei  $(V, \|\cdot\|)$  ein normierter Vektorraum. Dann gilt

$$| \|v\| - \|w\| | \leq \|v - w\| .$$

**Beweis.** Die Dreiecksungleichung zeigt

$$\|v\| = \|v - w + w\| \leq \|v - w\| + \|w\|, \quad \|w\| = \|w - v + v\| \leq \|v - w\| + \|v\|.$$

Zusammen zeigt dies die Ungleichung.  $\square$

**Beweis** von Satz 11.5. Sei  $\mathcal{B} = (b_1, \dots, b_N)$  eine geordnete Basis von  $V$ . Dann definieren wir zwei Funktionen  $f, f_\sim : \mathbb{K}^N \rightarrow \mathbb{R}_{\geq 0}$  durch

$$f(x) = \|(I_{\mathcal{B}})^{-1}(x)\|, \quad f_\sim(x) = \|(I_{\mathcal{B}})^{-1}(x)\|_\sim, \quad x \in \mathbb{K}^N.$$

Nun sind diese beiden Funktionen stetig, was z.B. für  $f$  gezeigt wird. Nach Lemma 11.6

$$\begin{aligned} |f(x) - f(y)| &= \left| \|(I_{\mathcal{B}})^{-1}(x)\| - \|(I_{\mathcal{B}})^{-1}(y)\| \right| \leq \|(I_{\mathcal{B}})^{-1}(x) - (I_{\mathcal{B}})^{-1}(y)\| \\ &= \left\| \sum_{n=1}^N (x_n - y_n) b_n \right\| \leq \sum_{n=1}^N |x_n - y_n| \|b_n\| \leq \left( \sum_{n=1}^N |x_n - y_n|^2 \right)^{\frac{1}{2}} \left( \sum_{n=1}^N \|b_n\|^2 \right)^{\frac{1}{2}}, \end{aligned}$$

wobei im letzten Schritt die Cauchy-Schwarz Ungleichung verwandt wurde. Wenn nun  $\|\cdot\|_2$  die euklidische Norm im  $\mathbb{K}^N$  bezeichnet (d.h. die vom Skalarprodukt herstammende), so zeigt Obiges also

$$|f(x) - f(y)| \leq C \|x - y\|_2, \quad C = \left( \sum_{n=1}^N \|b_n\|^2 \right)^{\frac{1}{2}}.$$

Also ist  $f$  Lipschitz stetig, und  $f_\sim$  ebenso. Nun verwenden wir folgenden Sachverhalt, der in der Analysisvorlesung bewiesen wird: Eine stetige Funktion auf einer kompakten Menge nimmt ihr Maximum und ihr Minimum an. (Es sei daran erinnert, dass nach dem Satz von Heine-Borel die Kompaktheit einer Menge im  $\mathbb{K}^N$  dazu äquivalent ist, dass die Menge eine abgeschlossene und beschränkte Teilmenge ist.) Nun ist die  $(N - 1)$ -dimensionale Sphäre  $\mathbb{S}^{N-1} = \{x \in \mathbb{K}^N \mid \|x\|_2 = 1\}$  eine kompakte Menge, und somit sind folgende Größen wohldefiniert:

$$M = \max_{\|x\|_2=1} f_\sim(x), \quad m = \min_{\|x\|_2=1} f(x).$$

Also folgt

$$\|v\|_\sim = f_\sim(I_{\mathcal{B}}v) = \|I_{\mathcal{B}}v\|_2 f_\sim\left(\frac{I_{\mathcal{B}}v}{\|I_{\mathcal{B}}v\|_2}\right) \leq \|I_{\mathcal{B}}v\| M \leq \|I_{\mathcal{B}}v\|_2 \frac{M}{m} f\left(\frac{I_{\mathcal{B}}v}{\|I_{\mathcal{B}}v\|_2}\right) = \frac{M}{m} \|v\|.$$

Dies zeigt eine Ungleichung, die andere folgt durch Vertauschen der Normen.  $\square$

**11.7 Korollar** Sei  $V$  ein endlich dimensionaler Vektorraum,  $\mathcal{B} = (b_1, \dots, b_N)$  eine beliebige geordnete Basis von  $V$  und  $\|\cdot\|$  eine beliebige Norm. Dann konvergiert eine Folge  $(v_k)_{k \geq 1}$  von Vektoren  $v_k \in V$  gegen  $v$  bez. der Norm  $\|\cdot\|$  genau dann, wenn jede Komponente der Koordinatenabbildung in  $\mathbb{K}$  konvergiert, d.h.

$$\lim_{k \rightarrow \infty} I_{\mathcal{B}}(v_k)_n = I_{\mathcal{B}}(v)_n, \quad n = 1, \dots, N.$$

**Beweis.** Die Konvergenz  $v_k \rightarrow v$  bedeutet ja, dass  $\|v_k - v\| \rightarrow 0$ . Nun gilt aber nach obigem Beweis und (11.1)

$$\begin{aligned} \|v_k - v\| &\leq M \|I_{\mathcal{B}}(v_k - v)\|_2 \leq MN^{\frac{1}{2}} \|I_{\mathcal{B}}(v_k) - I_{\mathcal{B}}(v)\|_\infty \\ &\leq MN^{\frac{1}{2}} \|I_{\mathcal{B}}(v_k) - I_{\mathcal{B}}(v)\|_2 \leq \frac{MN^{\frac{1}{2}}}{m} \|v_k - v\|. \end{aligned}$$

Also gehen alle Folgen gleichzeitig gegen 0 im Limes  $k \rightarrow \infty$ . Aber Konvergenz von  $\|I_B(v_k) - I_B(v)\|_\infty$  bedeutet gerade komponentenweise Konvergenz.  $\square$

In diesem Kapitel soll es um Normen auf dem Raum der linearen Abbildungen und Matrizen gehen. Eine abstrakte Klasse solcher Normen sind in folgendem Satz gegeben.

**11.8 Definition** Seien  $(V, \|\cdot\|_V)$  und  $(W, \|\cdot\|_W)$  normierte Vektorräume über  $\mathbb{K}$ . Die Operatornorm einer linearen Abbildung  $T \in \mathcal{L}(V, W)$  ist dann definiert als

$$\|T\|_{V \rightarrow W} = \sup_{v \in V \setminus \{0\}} \frac{\|Tv\|_W}{\|v\|_V} = \sup_{v \in V, \|v\|_V=1} \|Tv\|_W.$$

Falls  $W = V$ , wird die Operatornorm auch einfach mit  $\|T\|_V$  bezeichnet.

**11.9 Satz** Die Operatornorm  $\|\cdot\|_{V \rightarrow W}$  ist eine Norm auf dem Vektorraum  $\mathcal{L}(V, W)$ . Falls  $W = V$ , erfüllt sie außerdem

$$\|TS\|_V \leq \|T\|_V \|S\|_V, \quad T, S \in \mathcal{L}(V).$$

**Beweis.** Wieder sind Homogenität und Nicht-Entartung leicht zu überprüfen. Die Dreiecksungleichung folgt aus

$$\begin{aligned} \|T + S\|_{V \rightarrow W} &= \sup_{v \in V \setminus \{0\}} \frac{\|(T + S)v\|_W}{\|v\|_V} \leq \sup_{v \in V \setminus \{0\}} \frac{\|Tv\|_W + \|Sv\|_W}{\|v\|_V} \\ &\leq \sup_{v \in V \setminus \{0\}} \frac{\|Tv\|_W}{\|v\|_V} + \sup_{v \in V \setminus \{0\}} \frac{\|Sv\|_W}{\|v\|_V} = \|T\|_{V \rightarrow W} + \|S\|_{V \rightarrow W}. \end{aligned}$$

Analog

$$\|TS\|_V = \sup_{v \in V \setminus \{0\}} \frac{\|TSv\|_W}{\|Sv\|_V} \frac{\|Sv\|_W}{\|v\|_V} \leq \sup_{v \in V \setminus \{0\}} \frac{\|TSv\|_W}{\|Sv\|_V} \sup_{v \in V \setminus \{0\}} \frac{\|Sv\|_W}{\|v\|_V} = \|T\|_V \|S\|_V,$$

was den Beweis beendet.  $\square$

**11.10 Definition** Eine Norm  $\|\cdot\|$  auf dem Vektorraum  $\text{Mat}(N \times N, \mathbb{K})$  heißt Matrixnorm, wenn die Submultiplikativität

$$\|AB\| \leq \|A\| \|B\|$$

gilt.

Nach obigem Satz sind die Operatornormen allesamt Matrixnormen. Es gibt aber noch andere Matrixnormen, wie wir weiter unten sehen werden. Es gibt also sehr viele Normen und viele Gründe für diese Vielfalt (z.B. kann es einfach sein, Kontraktionseigenschaften bez. einer bestimmten Norm nachzuweisen, aber sehr schwierig oder sogar unmöglich bez. einer anderen). Weil es so viele Normen gibt, muss in einem mathematischen Text immer genau spezifiziert werden, welche Norm gerade verwandt wird. Unter all den Matrixnormen wird eine spezielle ganz besonders oft verwandt und ist in gewisser Weise auch besonders natürlich, und zwar die von dem euklidischen Skalarprodukt induzierte Operatornorm  $\|\cdot\|_{2 \rightarrow 2}$ . Hier wird sie bezeichnet mit

$$\|A\|_{\text{op}} = \|A\|_{2 \rightarrow 2} = \left( \sup_{v \in \mathbb{K}^N \setminus \{0\}} \frac{\langle Av | Av \rangle}{\langle v | v \rangle} \right)^{\frac{1}{2}} = \sup_{v \in \mathbb{K}^N, \|v\|=1} \langle Av | Av \rangle^{\frac{1}{2}}. \quad (11.2)$$

Wenn immer ein Vektorraum mit Skalarprodukt vorliegt (oder ein Hilbert-Raum), dann wird diese Norm schlichtweg als *die* Operatornorm bezeichnet. In dem folgenden Satz werden die wichtigsten Eigenschaften dieser Norm zusammengefasst.

**11.11 Satz** Es gilt für  $\|\cdot\|_{\text{op}}$  auf  $\text{Mat}(N \times N, \mathbb{K})$

$$\|A\|_{\text{op}} = \sup_{\|v\|=1} \sup_{\|w\|=1} |\langle v|Aw\rangle| = \max_{\|v\|=1} \max_{\|w\|=1} |\langle v|Aw\rangle|,$$

wobei  $\|v\| = \sqrt{\langle v|v\rangle}$ . Des Weiteren:

(i)  $\|A\|_{\text{op}} = \|A^*\|_{\text{op}}$

(ii) Die Operatornorm erfüllt die sogenannte  $C^*$ -Gleichung  $\|A^*A\|_{\text{op}} = \|A\|_{\text{op}}^2$ .

(iii)  $\|A\|_{\text{op}} = \max \{ |z|^{\frac{1}{2}} \mid z \in \text{Spec}(A^*A) \} = \mu_1$ , wobei  $\mu_1$  der größte Singulärwert von  $A$  ist.

(iv) Die Operatornorm ist unitär invariant, d.h.  $\|UAW\|_{\text{op}} = \|A\|_{\text{op}}$  für alle unitären Matrizen  $U$  und  $W$ .

(v) Falls  $A$  normal, so ist die Operatornorm gleich dem Spektralradius  $\|A\|_{\text{op}} = \max \{ |z| \mid z \in \text{Spec}(A) \}$ .

**Beweis.** Die erste Behauptung folgt direkt aus Skalierung und Korollar 7.14:

$$\|A\|_{\text{op}} = \sup_{\|v\|=1} \|Av\| = \sup_{\|v\|=1} \sup_{\|w\|=1} |\langle w|Av\rangle|.$$

Die Einheitskugel  $\{v \in \mathbb{K}^N \mid \|v\| = 1\}$  ist kompakt und somit werden die Suprema angenommen, d.h. die Maxima existieren. Nun zu (i). Nach Obigem

$$\|A\|_{\text{op}} = \sup_{\|v\|=1} \sup_{\|w\|=1} |\langle A^*w|v\rangle| = \sup_{\|w\|=1} \sup_{\|v\|=1} |\langle v|A^*w\rangle| = \|A^*\|_{\text{op}}.$$

Somit folgt auch

$$\|A^*A\|_{\text{op}} \leq \|A^*\|_{\text{op}} \|A\|_{\text{op}} = \|A\|_{\text{op}}^2,$$

also die erste Ungleichung in der  $C^*$ -Gleichung. Die Umkehrung kann wie folgt gezeigt werden:

$$\|A^*A\|_{\text{op}} = \sup_{\|v\|=1} \|A^*Av\| \geq \sup_{\|v\|=1} \langle v|A^*Av\rangle = \sup_{\|v\|=1} \langle Av|Av\rangle = \|A\|_{\text{op}}^2.$$

Zu (iii): Nach (11.2) gilt

$$\|A\|_{\text{op}} = \sup_{v \in \mathbb{K}^N, \|v\|=1} \langle Av|Av\rangle^{\frac{1}{2}} = \sup_{v \in \mathbb{K}^N, \|v\|=1} \langle v|A^*Av\rangle^{\frac{1}{2}} = (\mu_1)^{\frac{1}{2}},$$

wobei  $\mu_1$  der größte Singulärwert von  $A$  ist und die letzte Gleichheit aus dem Minimax-Prinzip folgt. Dies zeigt die Behauptung, denn die rechte Seite in (iii) ist genau  $(\mu_1)^{\frac{1}{2}}$ . Auch für (iv) beginnen wir von (11.2)

$$\|UAW\|_{\text{op}}^2 = \sup_{v \in \mathbb{K}^N \setminus \{0\}} \frac{\langle UAWv|UAWv\rangle}{\langle v|v\rangle} = \sup_{v \in \mathbb{K}^N \setminus \{0\}} \frac{\langle A(Wv)|A(Wv)\rangle}{\langle Wv|Wv\rangle} = \|A\|_{\text{op}}^2,$$

Letzteres weil  $W$  ja auch eine Bijektion auf  $\mathbb{K}^N \setminus \{0\}$  ist. Für den Beweis von (v) kann jetzt die unitäre Invarianz in Verbindung mit dem Spektralsatz für normale Matrizen verwandt werden. Sei also  $UAU^* = D$  diagonal mit den Eigenwerten  $z_1, \dots, z_N$  von  $A$  auf der Diagonalen, dann gilt

$$\|A\|_{\text{op}}^2 = \|D\|_{\text{op}}^2 = \sup_{\|v\|=1} \langle Dv|Dv\rangle = \sup_{\|v\|=1} \sum_{n=1}^N |v_n|^2 |z_n|^2 = \max_{n=1, \dots, N} |z_n|^2,$$

was genau die Behauptung ist. □

**11.12 Satz** Es sei  $f$  eine konvergente Potenzreihe

$$f(z) = \sum_{k \geq 0} f_k z^k, \quad f_k \in \mathbb{C},$$

mit Konvergenzradius  $\rho$ . Wenn  $A \in \text{Mat}(N \times N, \mathbb{K})$  eine Operatornorm hat mit  $\|A\|_{\text{op}} < \rho$ , dann konvergiert die Reihe

$$f(A) = \sum_{k \geq 0} f_k A^k.$$

**Beweis.** Da  $\|A^k\|_{\text{op}} \leq \|A\|_{\text{op}}^k$  nach der Submultiplikativität, ist die Reihe absolut konvergent, da

$$\sum_{k \geq 0} |f_k| \|A^k\|_{\text{op}} \leq \sum_{k \geq 0} |f_k| \|A\|_{\text{op}}^k \leq \sum_{k \geq 0} |f_k| (\rho - \epsilon)^k < \infty,$$

wobei  $\epsilon = \rho - \|A\|_{\text{op}} > 0$  nach Voraussetzung. □

**11.13 Beispiel** Da die Exponentialreihe überall auf  $\mathbb{C}$  konvergiert (Konvergenzradius  $\infty$ ), ist  $\exp(A)$  für jede Matrix wohldefiniert, ebenso  $\exp(tA)$  für jedes  $t$ . Ein anderes Beispiel ist die Neumann Reihe

$$(\mathbf{1} - A)^{-1} = \sum_{k \geq 0} A^k,$$

die konvergiert vorausgesetzt, dass  $\|A\|_{\text{op}} < 1$ . ◇

Nun soll eine andere wichtige Klasse von Normen betrachtet werden, die mit Hilfe der Spur definiert werden. Es sei daran erinnert (Satz 6.3), dass die Spur die Summe der Diagonaleinträge ist.

$$\text{Tr}(A) = \sum_{n=1}^N A_{n,n}.$$

Zunächst benötigen wir folgenden Satz, der elementare Eigenschaften der Spur zusammenfasst.

**11.14 Satz** Seien  $A, B \in \text{Mat}(N \times N, \mathbb{K})$  und  $\lambda \in \mathbb{K}$ , sowie  $U \in \text{Mat}(N \times N, \mathbb{K})$  unitär bzw. orthogonal. Dann

- (i)  $\text{Tr}(A + B) = \text{Tr}(A) + \text{Tr}(B)$
- (ii)  $\text{Tr}(\lambda A) = \lambda \text{Tr}(A)$
- (iii) (Zyklizität der Spur)  $\text{Tr}(AB) = \text{Tr}(BA)$
- (iv) (Unitäre Invarianz der Spur)  $\text{Tr}(A) = \text{Tr}(U^*AU)$
- (v) Für jede Orthonormalbasis  $u_1, \dots, u_N$  von  $\mathbb{K}^N$  gilt

$$\text{Tr}(A) = \sum_{n=1}^N \langle u_n | A | u_n \rangle.$$

- (vi) (Cauchy-Schwarz Ungleichung)  $|\text{Tr}(AB)| \leq \text{Tr}(A^*A)^{\frac{1}{2}} \text{Tr}(B^*B)^{\frac{1}{2}}$

(vii) (Standardabschätzung für die Spur)  $|\text{Tr}(AB)| \leq \|B\|_{\text{op}} \text{Tr}(|A|)$

**Beweis.** Die beiden ersten Punkte (i) und (ii) sind offensichtlich, und auch (iii) kann direkt nachgerechnet werden:

$$\text{Tr}(AB) = \sum_{n=1}^N (AB)_{n,n} = \sum_{n,m=1}^N A_{n,m} B_{m,n} = \sum_{m=1}^N (BA)_{m,m} = \text{Tr}(BA).$$

Hieraus folgt auch (iv), denn

$$\text{Tr}(U^*AU) = \text{Tr}(UU^*A) = \text{Tr}(A).$$

Zudem ist (v) lediglich ein Umschreiben von (iv). Zuletzt

$$|\text{Tr}(AB)| = \left| \sum_{n,m=1}^N A_{n,m} B_{m,n} \right| \leq \sum_{n,m=1}^N |A_{n,m}| |B_{m,n}| \leq \left( \sum_{n,m=1}^N |A_{n,m}|^2 \right)^{\frac{1}{2}} \left( \sum_{n,m=1}^N |B_{m,n}|^2 \right)^{\frac{1}{2}},$$

wobei im letzten Schritt die Cauchy-Schwarz Ungleichung für Vektoren der Länge  $N^2$  verwandt wurde. Die Identität

$$\sum_{n,m=1}^N |A_{n,m}|^2 = \text{Tr}(A^*A)$$

beendet den Beweis von (vi). Für den Beweis von (vii) wird die Polarzerlegung  $A = U|A|$  verwandt, sowie eine Orthonormalbasis  $u_1, \dots, u_N$  von Eigenvektoren von  $|A|$ :

$$\text{Tr}(AB) = \text{Tr}(U|A|B) = \text{Tr}(|A|BU) = \sum_{n=1}^N \mu_n(A) \langle u_n | BU u_n \rangle,$$

wobei  $\mu_n(A)$  die Singulärwerte von  $A$  sind, d.h. die Eigenwerte von  $|A|$ . Nun gilt  $|\langle u | Bv \rangle| \leq \|B\|_{\text{op}}$  für zwei normierte Vektoren  $u, v$ . Somit

$$|\text{Tr}(AB)| \leq \sum_{n=1}^N \mu_n(A) \|B\|_{\text{op}} = \text{Tr}(|A|) \|B\|_{\text{op}},$$

und der Beweis ist beendet. □

Des Weiteren sei daran erinnert, dass der Betrag einer Matrix definiert ist als  $|A| = (A^*A)^{\frac{1}{2}}$  durch Spektralkalkül der positiven Matrix  $A^*A$  (siehe Definition 9.15). Analog kann auch  $|A|^p = (A^*A)^{\frac{p}{2}}$  definiert werden.

**11.15 Definition** Sei  $p \geq 1$ . Die  $p$ -te Spurnorm ist definiert durch

$$\|A\|_p = \text{Tr}(|A|^p)^{\frac{1}{p}} \in [0, \infty).$$

Oft wird diese Norm auch als die  $p$ -te Schatten-Norm bezeichnet.

**11.16 Bemerkungen** 1. Selbstverständlich muss weiter unten erst nachgewiesen werden, dass die  $p$ -te Spurnorm tatsächlich eine Norm ist.



2. Die  $p$ -te Spurnorm kann berechnet werden durch die  $p$ -Norm des Vektors der Singulärwerte. In der Tat, sei  $U$  eine unitäre Matrix, die  $A^*A$  diagonalisiert, so dass auch

$$|A|^p = U^* \begin{pmatrix} (\mu_1)^p & & \\ & \ddots & \\ & & (\mu_N)^p \end{pmatrix} U,$$

wobei  $\mu_1, \dots, \mu_N$  die Singulärwerte von  $A$  sind. Somit gilt nach Satz 11.14(iv)

$$\|A\|_p = \text{Tr}(|A|^p)^{\frac{1}{p}} = \left( \sum_{n=1}^N (\mu_n)^p \right)^{\frac{1}{p}}. \quad (11.3)$$

3. Im Fall  $p = 2$  ist die Spurnorm explizit gegeben durch

$$\|A\|_2 = \left( \sum_{n=1}^N (A^*A)_{n,n} \right)^{\frac{1}{2}} = \left( \sum_{n,m=1}^N |A_{n,m}|^2 \right)^{\frac{1}{2}}.$$

Dies ist also genau die 2-Norm im Sinne des Beispiels 11.4 des  $N^2$ -komponentigen Vektors gegeben durch Untereinanderreihung der Spaltenvektoren von  $A$ . Die 2-te Spurnorm heisst auch die Frobenius-Norm oder die Hilbert-Schmidt Norm. Insbesondere ist für den Fall  $p = 2$  also auch schon nachgewiesen, dass  $\|\cdot\|_2$  eine Norm ist. Diese Norm ist von dem Skalarprodukt (Nachweis als Übung)

$$\langle A|B \rangle = \text{Tr}(A^*B),$$

auf  $\text{Mat}(N \times N, \mathbb{K})$  induziert. Die Ungleichung in Satz 11.14(vi) kann als die Cauchy-Schwarz-Ungleichung bez. dieses Skalarproduktes interpretiert werden, d.h.

$$|\langle A|B \rangle| \leq \|A\|_2 \|B\|_2.$$

4. Es sei darauf hingewiesen, dass für  $p \neq 2$  die Norm  $\|A\|_p$  nichts gemein hat mit der  $p$ -Norm aus Beispiel 11.4, obwohl die gleiche Notation dies nahelegen scheint. Es wäre zwar möglich, die Matrix  $A$  als einen Vektor der Länge  $N^2$  aufzufassen und davon die  $p$ -Norm zu betrachten, aber dies ist dann nicht die  $p$ -te Spurnorm.  $\diamond$

**11.17 Satz** Die  $p$ -te Spurnorm ist eine Norm für  $p \geq 1$ . Sie erfüllt des Weiteren:

- (i)  $\|A^*\|_p = \|A\|_p$
- (ii)  $\|AB\|_p \leq \|A\|_p \|B\|_{\text{op}}$  und  $\|AB\|_p \leq \|A\|_{\text{op}} \|B\|_p$
- (iii)  $\|A\|_{\text{op}} \leq \|A\|_p$
- (iv)  $\|AB\|_p \leq \|A\|_p \|B\|_p$  und somit ist die  $p$ -te Spurnorm eine Matrixnorm.

**Beweis.** Die Homogenität folgt aus

$$\|\lambda A\|_p = \text{Tr}(|\lambda A|^p)^{\frac{1}{p}} = \text{Tr}(|\lambda|^p |A|^p)^{\frac{1}{p}} = |\lambda| \text{Tr}(|A|^p)^{\frac{1}{p}} = |\lambda| \|A\|_p.$$

Die Nicht-Entartung folgt aus der Tatsache, dass lediglich für die Null-Matrix sämtliche Singulärwerte verschwinden. Schwierig ist hingegen der Nachweis der Dreiecksungleichung. Zunächst beginnen wir mit dem Fall  $p = 1$ , d.h. das Ziel ist es nachzuweisen, dass

$$\|A + B\|_1 \leq \|A\|_1 + \|B\|_1 .$$

Hierzu verwenden wir die Polarzerlegungen (Satz 9.20)

$$A + B = U |A + B| , \quad A = V |A| , \quad B = W |B| ,$$

wobei  $U, V, W$  unitär sind. Dann gilt

$$\|A + B\|_1 = \text{Tr}(|A + B|) = \text{Tr}(U^*(A + B)) = \text{Tr}(U^*V|A|) + \text{Tr}(U^*W|B|) .$$

Unter Verwendung der Cauchy-Schwarz Ungleichung (Satz 11.14 (vi)) folgt nun

$$\text{Tr}(U^*V|A|) = \text{Tr}(U^*V|A|^{\frac{1}{2}}|A|^{\frac{1}{2}}) \leq \text{Tr}\left(\left((U^*V|A|^{\frac{1}{2}})^*(U^*V|A|^{\frac{1}{2}})\right)^{\frac{1}{2}} \text{Tr}(|A|^{\frac{1}{2}})\right) = \text{Tr}(|A|) = \|A\|_1 .$$

Eine analoge Ungleichung für  $\text{Tr}(U^*W|B|)$  beendet den Beweis der Dreiecksungleichung für  $p = 1$ . Für den Fall  $p \neq 1, 2$  wird die folgende Ungleichung benötigt.  $\square$

**11.18 Satz (Hölder-von Neumann Ungleichung)** Seien  $p, q \geq 1$  so, dass  $\frac{1}{p} + \frac{1}{q} = 1$ . Dann gilt

$$\text{Tr}(|AB|) \leq \|A\|_p \|B\|_q .$$

**Beweis.** Wir betrachten den Fall  $p \leq 2$  und  $q \geq 2$ , der andere geht analog. Wieder beginnen wir mit den Polarzerlegungen

$$AB = U |AB| , \quad A = V |A| , \quad B = W |B| .$$

Dann gilt, wenn die Spur in der orthonormalen Eigenbasis  $u_1, \dots, u_N$  von  $|A|$  berechnet wird,

$$\text{Tr}(|AB|) = \text{Tr}(U^*V|A|W|B|) = \text{Tr}(|A|W|B|U^*V) = \sum_{n=1}^N \mu_n(A) \langle W^*u_n | |B| U^*V u_n \rangle .$$

Nun bilden auch  $v_n = W^*u_n, n = 1, \dots, N$  eine Orthonormalbasis, sowie  $w_n = U^*V u_n, n = 1, \dots, N$ . Für beliebige normierte Vektoren  $v$  und  $w$  gilt, wenn  $x_1, \dots, x_N$  eine Orthonormalbasis von Eigenvektoren von  $B^*B$  ist,

$$\langle v | |B| w \rangle \leq \left( \langle v | |B|^2 v \rangle \right)^{\frac{1}{2}} = \left( \sum_{n=1}^N |\langle v | x_n \rangle|^2 \mu_n(B)^2 \right)^{\frac{1}{2}} \leq \left( \sum_{n=1}^N |\langle v | x_n \rangle|^2 \mu_n(B)^q \right)^{\frac{1}{q}} ,$$

wobei die Jensen Ungleichung für die konvexe Funktion  $y \in \mathbb{R}_{\geq 0} \mapsto y^{\frac{q}{2}}$  verwandt wurde. Somit

$$\langle v | |B| w \rangle \leq \langle v | |B|^q v \rangle^{\frac{1}{q}} .$$

Wird dies nun oben für jedes  $n$  verwandt, so folgt

$$\text{Tr}(|AB|) \leq \sum_{n=1}^N \mu_n(A) \langle v_n | |B|^q v_n \rangle^{\frac{1}{q}} \leq \left( \sum_{n=1}^N \mu_n(A)^p \right)^{\frac{1}{p}} \left( \sum_{n=1}^N \langle v_n | |B|^q v_n \rangle \right)^{\frac{1}{q}} = \|A\|_p \|B\|_q ,$$

wobei im vorletzten Schritt die Hölder Ungleichung verwandt wurde. □

**Beweis** der Minkovski Ungleichung

$$\|A + B\|_p \leq \|A\|_p + \|B\|_p .$$

Wiederum werden die Polarzerlegungen verwandt

$$A + B = U |A + B| , \quad A = V |A| , \quad B = W |B| .$$

Jetzt

$$\begin{aligned} \text{Tr}(|A + B|^p) &= \text{Tr}(|A + B| |A + B|^{p-1}) = \text{Tr}(U^*(A + B) |A + B|^{p-1}) \\ &= \text{Tr}(U^*V|A| |A + B|^{p-1}) + \text{Tr}(U^*W|B| |A + B|^{p-1}) \\ &\leq \text{Tr}(|A| |A + B|^{p-1}) + \text{Tr}(|B| |A + B|^{p-1}) , \end{aligned}$$

Letzteres nach Anwendung von Satz 11.14(vii). Nun kann die Hölder Ungleichung aus Satz 11.18 zweimal angewandt werden:

$$\|A+B\|_p^p = \text{Tr}(|A+B|^p) \leq \|A\|_p \| |A+B|^{p-1} \|_q + \|B\|_p \| |A+B|^{p-1} \|_q = (\|A\|_p + \|B\|_p) \|A+B\|_p^{\frac{p}{q}} ,$$

wobei die Identität  $(p-1)q = p$  verwandt wurde. Da  $p - \frac{p}{q} = 1$ , folgt die Minkovski Ungleichung.

Es verbleibt noch, die Eigenschaften (i)-(iv) in Satz 11.17 zu nachzuweisen. (i) und (ii) folgen aus dem unten stehenden Lemma für die Singulärwerte kombiniert mit (11.3). (iii) folgt ebenso aus (11.3) und der Identität  $\mu_1(A) = \|A\|_{\text{op}}$ , bewiesen in Satz 11.11(iii) und (iv) ist dann eine Folgerung aus (ii) und (iii). □

**11.19 Lemma** Die Sinulärwerte  $\mu_n(A)$  von  $A$  erfüllen

$$\mu_n(A^*) = \mu_n(A) , \quad \mu_n(AB) \leq \|A\|_{\text{op}} \mu_n(B) , \quad \mu_n(AB) \leq \mu_n(A) \|B\|_{\text{op}} .$$

**Beweis.** Die erste Gleichheit folgt z.B. aus der Singulärwertzerlegung (Satz 9.18). Weiter gilt wegen  $A^*A \leq \|A\|_{\text{op}}^2 \mathbf{1}$  (was per Definition bedeutet  $\|A\|_{\text{op}}^2 \mathbf{1} - A^*A \geq 0$ ) und Korollar 9.10

$$(AB)^*(AB) = B^*A^*AB \leq \|A\|_{\text{op}}^2 B^*B .$$

Somit folgt aus dem Minimax-Prinzip  $\mu_n(AB) \leq \|A\|_{\text{op}} \mu_n(B)$ . Die letzte Ungleichung folgt nun aus  $\mu_n(AB) = \mu_n(B^*A^*) \leq \|B^*\|_{\text{op}} \mu_n(A^*) = \|B\|_{\text{op}} \mu_n(A)$ . □

## 12 Lineare Optimierung

Dieses Kapitel folgt sehr nah dem Buch von Schaback und Wendland [SW].

**12.1 Definition** Seien  $p \in \mathbb{R}^N$ ,  $b \in \mathbb{R}^M$ ,  $c \in \mathbb{R}^K$ ,  $A \in \text{Mat}(M \times N, \mathbb{R})$  und  $B \in \text{Mat}(K \times N, \mathbb{R})$ . Ein lineares Optimierungsproblem, auch lineares Programm genannt, besteht aus die Minimierung einer reellen linearen Kostenfunktion

$$x \in \mathbb{R}^N \mapsto \langle p|x \rangle \in \mathbb{R}$$

unter den linearen Nebenbedingungen

$$Ax = b , \quad Bx \leq c , \tag{12.1}$$

wobei die letzte Ungleichung komponentenweise zu verstehen ist, d.h.  $(Bx)_k \leq c_k$  für  $k = 1, \dots, K$ . Punkte  $x$ , die (12.1) erfüllen, heissen zulässig.

**12.2 Bemerkungen** 1. Es ist möglich nach einem Maximum zu suchen, indem  $p$  durch  $-p$  ersetzt wird.

2. Falls eher eine umgekehrte Ungleichung  $(Bx)_k \geq c_k$  vorliegt, so kann die  $k$ -te Zeile von  $B$  mit  $-1$  multipliziert werden, so dass für das modifizierte  $B$  wieder die Ungleichung  $(Bx)_k \leq c_k$  gilt.

3. Wir werden sehen, dass das Optimierungsproblem keine Lösung haben kann, oder genau eine, oder (sogar unendlich) viele.  $\diamond$

Das Problem soll zuerst auf eine Normalform gebracht werden. Zunächst fügen wir sogenannte Schlupfvariablen  $y_1, \dots, y_K \geq 0$  ein, die einen Vektor  $y \in \mathbb{R}^K$  mit  $y \geq 0$  bilden. Dann kann die Ungleichung  $Bx \leq c$  geschrieben werden als eine Gleichung zusammen mit einer Positivitätsbedingung

$$(B \mathbf{1}_K) \begin{pmatrix} x \\ y \end{pmatrix} = c, \quad y \geq 0.$$

Nun wird  $p$  durch Nulleinträge ergänzt zu einem Vektor in  $\mathbb{R}^{N+K}$  und  $A$  durch Nullspalten zu einer Matrix der Größe  $M \times (N + K)$ . Als Nächstes sollen auch die Komponenten von  $x$  noch ein definites Vorzeichen erhalten. Dafür setzen wir

$$x^\pm = \begin{pmatrix} x_1^\pm \\ \vdots \\ x_N^\pm \end{pmatrix}, \quad x_n^+ = \max\{0, x_n\}, \quad x_n^- = \max\{0, -x_n\}.$$

Also  $x = x^+ - x^-$ . Mit diesen Definitionen kann die Kostenfunktion geschrieben werden als

$$\begin{pmatrix} x^+ \\ x^- \\ y \end{pmatrix} \in \mathbb{R}^{2N+K} \mapsto \left\langle \begin{pmatrix} p \\ -p \\ 0 \end{pmatrix} \mid \begin{pmatrix} x^+ \\ x^- \\ y \end{pmatrix} \right\rangle,$$

und die Nebenbedingungen als

$$\begin{pmatrix} A & -A & 0 \\ B & -B & \mathbf{1}_K \end{pmatrix} \begin{pmatrix} x^+ \\ x^- \\ y \end{pmatrix} = \begin{pmatrix} b \\ c \end{pmatrix}, \quad \begin{pmatrix} x^+ \\ x^- \\ y \end{pmatrix} \geq 0.$$

Es sei bemerkt, dass die zusätzlichen Variablen  $y$  nicht zu zusätzlichen Kosten führen. Also haben wir den folgenden Sachverhalt bewiesen:

**12.3 Satz** Jedes lineare Optimierungsproblem lässt sich in die folgende Normalform bringen: Das zu minimierende Kostenfunktional ist von der Form

$$x \in \mathbb{R}^N \mapsto \langle p \mid x \rangle \in \mathbb{R}$$

zugehörig zu  $p \in \mathbb{R}^N$ , und die Nebenbedingungen sind

$$Ax = b, \quad x \geq 0,$$

wobei  $A \in \text{Mat}(M \times N, \mathbb{R})$  ist und  $b \in \mathbb{R}^M$ .

Von nun an werden wir lediglich die Normalform betrachten. Dies ist in der mathematischen Darstellung etwas einfacher, auch wenn das in einem Programm wegen zusätzlicher Variablen nicht so implementiert wird (es gibt ohnehin sehr gute Software Pakete, die die lineare Optimierung implementieren, so dass es im Wesentlichen darum geht zu verstehen, was diese Programme prinzipiell machen). Zugehörig zur Normalform betrachten wir nun die Menge der zulässigen Punkte

$$P = \{x \in \mathbb{R}^N \mid Ax = b \text{ und } x \geq 0\} . \tag{12.2}$$

Geometrisch ist diese Menge vorstellbar als der Schnitt einer durch  $Ax = b$  gegebenen Hyperebene mit dem Orthanten  $x \geq 0$  (in Dimension  $N = 2$  Quadrant, in Dimension  $N = 3$  Oktant). Natürlich kann dieser Schnitt leer sein und in diesem Fall gibt es keine zulässigen Punkte und somit auch keine Lösung des Optimierungsproblems. Falls  $P$  nicht leer und beschränkt ist (und somit auch kompakt als abgeschlossene Menge), so nimmt die (stetige) Kostenfunktion auf jeden Fall ihr Minimum an und somit existiert eine Lösung des Optimierungsproblems. Wenn hingegen  $P$  unbeschränkt ist, ist es möglich, dass es kein Minimum auf  $P$  gibt (d.h. die Werte der Kostenfunktion werden beliebig klein). Im Folgenden werden zunächst eine Reihe elementarer Eigenschaften der zulässigen Menge bewiesen. Eine erste elementare Eigenschaft von  $P$  ist die folgende.

**12.4 Satz**  $P$  ist eine konvexe Teilmenge des  $\mathbb{R}^N$ , d.h. mit  $x, y \in P$  liegt auch die Verbindungsstrecke

$$[x, y] = \{\lambda x + (1 - \lambda)y \mid \lambda \in [0, 1]\}$$

ganz in  $P$ .

**Beweis.** Dies ist offensichtlich, da  $A(\lambda x + (1 - \lambda)y) = \lambda Ax + (1 - \lambda)Ay = \lambda b + (1 - \lambda)b = b$  und  $\lambda x + (1 - \lambda)y \geq 0$  falls  $x \geq 0$  und  $y \geq 0$ . □

**12.5 Definition** Ein Punkt  $x \in P$  heißt Ecke einer konvexen Menge  $P$  genau dann, wenn aus  $x \in [y, z]$  für  $y, z \in P$  folgt  $x = y$  oder  $x = z$ , d.h.  $x$  liegt nicht im Inneren einer Verbindungsstrecke in  $P$ .

Zugehörig zu  $x \in \mathbb{R}^N$  definieren wir die Indexmenge

$$I_x = \{n \in \{1, \dots, N\} \mid x_n \neq 0\} ,$$

sowie die nach  $I_x$  selektierten Untermatrizen von  $A = (a_1, \dots, a_N)$  als

$$A_x = (a_n)_{n \in I_x} \in \text{Mat}(M \times |I_x|, \mathbb{R}) ,$$

wobei  $|I_x|$  die Anzahl der Punkte in  $I_x$  bezeichnet. Entsprechend definieren wir für einen Vektor  $v \in \mathbb{R}^N$

$$v_x = (v_n)_{n \in I_x} \in \mathbb{R}^{|I_x|} .$$

Mit diesen Definitionen gilt dann

$$Ax = A_x x_x .$$

Etwas allgemeiner gilt für  $v \in \mathbb{R}^N$  mit  $I_v \subset I_x$  auch  $Av = A_x v_x$ .

**12.6 Satz** Sei  $x$  ein Punkt in der zulässigen Menge  $P$  definiert in (12.2). Dann sind folgende Bedingungen äquivalent:

- (i)  $x$  ist Ecke von  $P$ .

(ii) Die Spalten  $\{a_n \mid n \in I_x\}$  von  $A_x$  sind linear unabhängig.

(iii)  $\text{rk}(A_x) = |I_x|$ .

**Beweis.** (i) $\implies$ (ii): Seien die Spalten von  $A_x$  linear abhängig. Dann existiert ein  $y' \in \mathbb{R}^{|I_x|}$ , so dass  $A_x y' = 0$ . Nun wird  $y'$  durch Nullen ergänzt zu einem Vektor  $y \in \mathbb{R}^N$ . Es gilt  $I_{y'} = I_y \subset I_x$ . Somit gilt für hinreichend kleines  $\epsilon > 0$ , dass

$$x_\epsilon^\pm = x \pm \epsilon y \geq 0,$$

sowie

$$A x_\epsilon^\pm = A x \pm \epsilon A y = b \pm \epsilon A_x y' = b.$$

Also sind  $x_\epsilon^\pm \in P$  und  $x = \frac{1}{2}(x_\epsilon^+ - x_\epsilon^-) \in [x_\epsilon^-, x_\epsilon^+]$ , so dass  $x$  keine Ecke ist.

(ii) $\implies$ (i): Sei  $x$  keine Ecke. Somit existieren verschiedene Punkte  $y, z \in P$  und  $\lambda \in (0, 1)$  mit  $x = \lambda y + (1 - \lambda)z$ . Hieraus folgt  $I_y \subset I_x$  und  $I_z \subset I_x$ , denn aus  $x_n = 0$  folgt ja  $y_n = z_n = 0$ . Somit gilt auch  $I_{y-z} \subset I_x$  und deswegen

$$A_x(y - z)_x = A(y - z) = b - b = 0.$$

Nun ist aber  $(y - z)_x \neq 0$ , denn sonst wäre wegen  $I_{y-z} \subset I_x$  auch  $y = z$ . Somit folgt, dass die Zeilen von  $A_x$  linear abhängig sind.

(ii) $\iff$ (iii): Dies ist eine Umformulierung, die für jede Matrix mit linear unabhängigen Spalten gilt.  $\square$

**12.7 Satz** (i) Jede Ecke  $x$  ist durch ihre Indexmenge  $I_x$  eindeutig bestimmt.

(ii) Die Menge  $P$  definiert in (12.2) hat endlich viele Ecken.

(iii) Die Menge  $P$  definiert in (12.2) hat mindestens eine Ecke.

(iv) Falls das Optimierungsproblem lösbar ist, so wird das Minimum auch an einer Ecke angenommen.

**Beweis.** (i) In der Tat, die Ecke  $x$  erfüllt das Gleichungssystem  $Ax = A_x x_x = b$  und da der Rang von  $A_x$  nach Satz 12.6 gleich  $|I_x|$  ist und  $x_x \in \mathbb{R}^{|I_x|}$ , gibt es eine eindeutige Lösung  $x_x$ , die dann ja auch  $x$  festlegt, weil alle anderen Komponenten von  $x$  verschwinden. (ii) folgt aus (i), da es nur endlich viele Möglichkeiten gibt, aus den  $N$  Spalten von  $A$  linear unabhängige Spalten auszuwählen. (iii) Sei  $x \in P$  so gewählt, dass die Kardinalität der Indexmenge  $|I_x|$  minimal ist (es gibt immer solche Minimierer). Wir nehmen an, dass  $x$  keine Ecke ist und führen dies zu einem Widerspruch. Da  $x$  keine Ecke ist, gibt es  $y, z \in P$  mit  $x = \lambda y + (1 - \lambda)z$  für ein  $\lambda \in (0, 1)$ . Wie im Beweis von Satz 12.6 impliziert dies  $I_y \subset I_x$  und  $I_z \subset I_x$  sowie  $A_x(y - z)_x = 0$  mit  $(y - z)_x \neq 0$ . Es möge ein Komponentenindex  $n \in I_x$  existieren mit  $(y - z)_n < 0$  (wenn nur positive Einträge vorhanden sind, verwende  $z - y$  anstelle von  $y - z$ ). Wieder betrachten wir  $x + \epsilon(y - z) \in P$ , allerdings wird  $\epsilon > 0$  speziell gewählt als

$$\epsilon = \min \left\{ -\frac{x_n}{y_n - z_n} \mid n \in I_x \text{ mit } y_n - z_n < 0 \right\}.$$

Dieses Minimum existiert tatsächlich nach unserer Annahme. Wenn  $\tilde{n}$  einer der Indizes ist, bei denen das Minimum angenommen wird, dann gilt  $x_{\tilde{n}} + \epsilon(y - z)_{\tilde{n}} = 0$  und nach wie vor gilt  $u = x + \epsilon(y - z) \in P$ , weil wegen der Minimalität bei keinem anderen Index  $n$  das Vorzeichen von  $x_n + \epsilon(y - z)_n$  wechselt. Aber nun gilt nach Konstruktion  $|I_u| < |I_x|$  im Widerspruch zur Annahme, dass  $I_x$  minimal ist. Also ist  $x$  eine Ecke.

(iv) Innerhalb der Menge aller minimierenden Punkte  $x \in P$  mit

$$\langle p|x \rangle \leq \langle p|y \rangle \quad \text{für alle } y \in P,$$

sei wieder einer ausgewählt mit minimalen  $|I_x|$ . Wiederum zeigen wir, dass dieser Punkt  $x$  eine Ecke sein muss (dies folgt nicht aus obigem Argument, da  $|I_x|$  nicht globales Minimum sein muss, sondern lediglich die Indexmenge unter allen minimierenden Punkten minimiert). Ansonsten könnte man nämlich wie in (iii) ein  $u = x + \epsilon(y - z) \in P$  mit  $|I_u| < |I_x|$  konstruieren. Es ist dann auch  $v = x - \delta(y - z) \in P$  für  $\delta > 0$  ausreichend klein. Somit gilt

$$\frac{\delta}{\epsilon + \delta} u = \frac{\delta}{\epsilon + \delta} x + \frac{\delta\epsilon}{\epsilon + \delta} (y - z), \quad \frac{\epsilon}{\epsilon + \delta} v = \frac{\epsilon}{\epsilon + \delta} x - \frac{\delta\epsilon}{\epsilon + \delta} (y - z).$$

Mit  $\lambda = \frac{\delta}{\epsilon + \delta} \in (0, 1)$  gilt dann nach Summation

$$x = \lambda u + (1 - \lambda) v.$$

Dies impliziert

$$\langle p|x \rangle = \lambda \langle p|u \rangle + (1 - \lambda) \langle p|v \rangle,$$

d.h.  $\langle p|x \rangle$  liegt zwischen  $\langle p|u \rangle$  und  $\langle p|v \rangle$ . Somit müssen alle Werte gleich dem Minimum sein. Somit gilt der Minimalwert auch bei  $u$  mit  $|I_u| < |I_x|$  angenommen, im Widerspruch zur Annahme.  $\square$

Der folgende Satz liefert noch weitere geometrische Informationen über die Menge  $P$ .

**12.8 Satz** Wenn die Menge  $P$  definiert in (12.2) beschränkt ist, so ist sie ein konvexer Polyeder, d.h. sie ist die Konvexkombination (auch konvexe Hülle genannt) von seinen Ecken  $y_1, \dots, y_K \in \mathbb{R}^N$ :

$$P = \left\{ \sum_{k=1}^K \lambda_k y_k \mid \lambda_k \in [0, 1] \text{ mit } \sum_{k=1}^K \lambda_k = 1 \right\}.$$

Falls die Ecken noch zusätzlich affin unabhängig sind (d.h. die Vektoren  $y_1 - y_2, \dots, y_1 - y_K$  sind linear unabhängig), so spricht man auch von einem  $K$ -Simplex.

**Beweis.** Nach Satz 12.4 liegt das konvexe Polytop sicher in  $P$ , da ja auch die Ecken in  $P$  liegen. Es verbleibt also zu zeigen, dass jeder Punkt  $x \in P$  als Konvexkombination der Ecken geschrieben werden kann. Dies zeigen wir induktiv über  $r = |I_x|$ . Wenn  $r = 0$ , so ist (was nur vorkommt, wenn  $x = 0 \in P$ ), so ist  $x$  eine Ecke. Sei nun  $r > 0$ . Falls die  $r$  Spaltenvektoren von  $A_x$  linear unabhängig sind, ist  $x$  nach Satz 12.6 eine Ecke und somit nichts nachzuweisen. Anderenfalls sind die Spaltenvektoren von  $A_x$  linear abhängig und es ist, nach ähnlichen Argumenten wie oben, möglich  $x$  als inneren Punkt einer Verbindungsstrecke  $[y, z]$  mit  $y, z \in P$  zu schreiben. Wenn nun die Komponenten des Vektors  $y - z$  alle das gleiche Vorzeichen hätten, sagen wir nicht negativ, dann wäre die unbeschränkte Menge  $\{x + t(y - z) \mid t \geq 0\}$  in  $P$ , im Widerspruch zur Annahme, dass  $P$  beschränkt ist. Also hat  $y - z$  sowohl positive als auch negative Einträge und es ist möglich (wie im Beweis von Satz 12.7)  $\epsilon, \delta > 0$  zu bestimmen, so dass  $u = x + \epsilon(y - z) \in P$  und  $v = x - \delta(y - z) \in P$  mit  $|I_u| < |I_x|$  und  $|I_v| < |I_x|$ . Nun sind aber nach Induktionsvoraussetzung  $u$  und  $v$  Konvexkombination von Ecken von  $P$ , und somit ist es möglich auch  $x$  als Konvexkombination von Ecken zu schreiben (wiederum analog wie im Beweis von Satz 12.7).  $\square$

Satz 12.7 impliziert, dass im Falle einer kompakten Menge  $P$ , das Minimum bestimmt werden kann, indem einfach die Werte an sämtlichen Ecken berechnet werden und von diesen endlich vielen Werten der kleinste ausgewählt wird. In der Praxis kann es aber sehr viele Ecken geben, und dann ist dieses Verfahren nicht besonders effizient. Das von Dantzig 1948 vorgeschlagene sogenannte Simplex-Verfahren beginnt mit einer Ecke und sucht sich dann unter den benachbarten Ecken diejenige aus, die den kleinsten Wert hat. Dieser wird dann als nächste Ausgangsecke verwandt, und es wird wieder unter den benachbarten Ecken die mit dem kleinsten Kostenwert ausgewählt. Hierbei ist eine benachbarte Ecke eine solche, bei der aus der

Indexmenge  $I_x$  (die die Ecke  $x$  nach Satz 12.7 ja spezifiziert) einer der Indizes ausgewechselt wird (dabei muss natürlich die lineare Unabhängigkeit erhalten bleiben). Dieses Verfahren führt recht schnell zu der Ecke mit dem kleinsten Kostenwert. Dies soll im Folgenden etwas detaillierter beschrieben werden, unter folgende in Anwendungen meist erfüllter

**Annahme** Es ist  $N > M$  und der Rang von  $A \in \text{Mat}(M \times N, \mathbb{R})$  ist gleich  $M$ .

Zunächst ist klar, dass es notwendig ist, bestimmte Ecken zu unterscheiden.

**12.9 Definition** Eine Ecke  $x$  der zulässigen Menge  $P$  definiert in (12.2) heißt entartet genau dann, wenn  $|I_x| < M$ . Ansonsten heißt sie nicht entartet.

Wenn eine Ecke entartet ist, wird die Indexmenge  $I_x$  erweitert zu  $X = \{k_1, \dots, k_M\}$  dergestalt, dass  $A_X = (a_{k_1}, \dots, a_{k_M})$  maximalen Rang  $M$  hat. Dann ist  $A_X$  invertierbar. Hierbei ist  $X$  eine geordnete Menge durch die Folge  $l \in \{1, \dots, M\} \mapsto k_l$ , allerdings gilt nicht  $k_l < k_{l+1}$ . Der Übergang zu einer benachbarten Ecke ist nun gegeben durch einen

**Indexwechsel:** Wähle  $m \in \{1, \dots, M\}$  und  $n \in \{1, \dots, N\}$ , so dass das Ersetzen von  $k_m$  durch  $n$  zu eine Indexmenge  $Y = \{n\} \cup X \setminus k_m$  führt, für die  $A_Y = (a_{k_1}, \dots, a_{k_{m-1}}, a_n, a_{k_{m+1}}, \dots, a_{k_M})$  wieder Rang  $M$  hat.

Dann ist die Ecke  $y$  festgelegt durch  $A_Y y_Y = b$  und  $I_y \subset Y$  (Letzteres bedeutet, dass  $y$  durch Nullen außerhalb der Indexmenge  $Y$  aufgefüllt wird). Nun soll die Wahl von  $m$  und  $n$  geschickt durchgeführt werden und die neue Ecke  $y$  aus  $x$  sowie die Kosten bei  $y$  aus denen bei  $x$  berechnet werden, und dies zudem auf eine algorithmische Art und Weise, die den Rechenaufwand minimiert. Eine Möglichkeit dies zu erreichen ist die Folgende. Alle Daten des Problems werden in eine Matrix

$$\hat{A} = \begin{pmatrix} A & b \\ p^t & 0 \end{pmatrix} \in \text{Mat}((M+1) \times (N+1), \mathbb{R})$$

zusammengefasst, und zugehörig zu einer Indexmenge  $X$  wird gesetzt

$$B_X = \begin{pmatrix} A_X & 0 \\ (p_X)^t & -1 \end{pmatrix} \in \text{Mat}((M+1) \times (M+1), \mathbb{R})$$

Nun ist  $B_X$  offensichtlich invertierbar und deswegen existiert eine Matrix  $C_X \in \text{Mat}((M+1) \times (N+1), \mathbb{R})$  mit

$$\hat{A} = B_X C_X .$$

Das Inverse von  $B_X$  kann mit der Schurkomplement-Formel berechnet werden (oder einfach verifiziert werden), so dass

$$\begin{aligned} C_X &= (B_X)^{-1} \hat{A} = \begin{pmatrix} (A_X)^{-1} & 0 \\ (p_X)^t (A_X)^{-1} & -1 \end{pmatrix} \begin{pmatrix} A & b \\ p^t & 0 \end{pmatrix} = \begin{pmatrix} (A_X)^{-1} A & (A_X)^{-1} b \\ (p_X)^t (A_X)^{-1} A - p^t & (p_X)^t (A_X)^{-1} b \end{pmatrix} \\ &= \begin{pmatrix} (A_X)^{-1} A & x_X \\ (p_X)^t (A_X)^{-1} A - p^t & (p_X)^t x_X \end{pmatrix} = \begin{pmatrix} (A_X)^{-1} A & x_X \\ (p_X)^t (A_X)^{-1} A - p^t & p^t x \end{pmatrix} , \end{aligned}$$

wobei im letzten Schritt lediglich  $(p_X)^t x_X = p^t x$  verwandt wurde. Somit erhält  $C_X$  in der letzten Spalte die wichtigsten Informationen, nämlich die Ecke (in Form von  $x_X$ ) sowie die Kosten bei der Ecke (in Form von  $p^t x$ ). Genau die gleiche Faktorisierung von  $\hat{A}$  gilt auch für  $Y$ :

$$\hat{A} = B_X C_X = B_Y C_Y ,$$



und da sowohl  $B_X$  als auch  $B_Y$  invertierbar sind gibt es auch ein invertierbares

$$D = (B_Y)^{-1} B_X = \begin{pmatrix} (A_Y)^{-1} & 0 \\ (p_Y)^t (A_Y)^{-1} & -1 \end{pmatrix} \begin{pmatrix} A_X & 0 \\ (p_X)^t & -1 \end{pmatrix} = \begin{pmatrix} (A_Y)^{-1} A_X & 0 \\ (p_Y)^t (A_Y)^{-1} A_X - (p_X)^t & 1 \end{pmatrix}.$$

Nun unterscheiden sich  $A_X$  und  $A_Y$  ja lediglich durch die  $m$ -te Spalte und deswegen

$$(A_Y)^{-1} A_X = \begin{pmatrix} \mathbf{1}_{m-1} & & \\ & d & \\ & & \mathbf{1}_{M-m-1} \end{pmatrix},$$

mit einem Vektor  $d = (A_Y)^{-1} A_X e_m \in \mathbb{R}^M$ . Also, mit einem Vektor  $d' = \begin{pmatrix} d \\ d_{M+1} \end{pmatrix} = \begin{pmatrix} d \\ (p_Y)^t d - p_{k_m} \end{pmatrix} \in \mathbb{R}^{M+1}$ ,

$$D = \begin{pmatrix} \mathbf{1}_{m-1} & & \\ & d' & \\ & & \mathbf{1}_{M-m} \end{pmatrix}.$$

Es gibt jedoch eine alternative Art und Weise,  $d'$  und  $D$  zu berechnen. Die neue  $m$ -te Spalte von  $B_Y$  ist

$$B_Y e_m = \widehat{A} e_n = B_X C_X e_n = B_Y D C_X e_n.$$

Also ist wegen der Invertierbarkeit von  $B_Y$

$$e_m = D C_X e_n.$$

Setzen wir also

$$c = C_X e_n = \begin{pmatrix} c_1 \\ \vdots \\ c_{M+1} \end{pmatrix} \in \mathbb{R}^{M+1}, \quad (12.3)$$

so ergibt sich

$$d' = \begin{pmatrix} d_1 \\ \vdots \\ d_{M+1} \end{pmatrix} = \frac{1}{c_m} \begin{pmatrix} -c_1 \\ \vdots \\ -c_{m-1} \\ 1 \\ -c_{m+1} \\ \dots \\ c_{M+1} \end{pmatrix}.$$

Nun ist es möglich, die neuen Ecke und die neuen Kosten aus den alten zu berechnen:

$$\begin{pmatrix} y_Y \\ p^t y \end{pmatrix} = C_Y e_{M+1} = (B_Y)^{-1} B_X C_X e_{M+1} = D \begin{pmatrix} x_X \\ p^t x \end{pmatrix}.$$

Expliziter ausgeschrieben ist dies für die neuen Komponenten des Eckenvektors  $y$ :

$$y_{k_l} = x_{k_l} + d_l x_{k_m} = x_{k_l} - \frac{c_l}{c_m} x_{k_m} \quad \text{für } l \neq m, \quad \text{und} \quad y_n = d_m x_{k_m} = \frac{1}{c_m} x_{k_m}. \quad (12.4)$$

Selbstverständlich muss dieser Vektor  $y \geq 0$  erfüllen. Somit ist es notwendig,  $m$  so zu wählen, dass  $c_m > 0$ . Weiter unten in Satz 12.10 wird gezeigt, dass es keine Lösung gibt, falls  $c_l < 0$  für alle  $l = 1, \dots, M$ .

Außerdem muss für all die  $l$  mit  $c_l > 0$  wegen der Formel für  $y_{k_l}$  auch gelten  $\frac{x_{k_l}}{c_l} - \frac{x_{k_m}}{c_m} \geq 0$ . Deswegen muss  $m$  so gewählt werden, dass

$$\frac{x_{k_m}}{c_m} = \min \left\{ \frac{x_{k_l}}{c_l} \mid l = 1, \dots, M, c_l > 0 \right\}. \quad (12.5)$$

Also ist  $m$  festgelegt, wenn schon  $n$  gewählt wurde, denn  $n$  legt wegen (12.3) auch  $c$  und somit alles Weitere fest. Für die richtige Wahl von  $n$  wird nun die Kostenfunktion bei  $y$  mit der bei  $x$  verglichen. Einsetzen von (12.4) liefert zunächst

$$\langle p|y \rangle = p^t y = p_n \frac{x_{k_m}}{c_m} + \sum_{l=1}^M p_{k_l} \left( x_{k_l} - \frac{c_l}{c_m} x_{k_m} \right),$$

wobei in der Summe der Summand  $l = m$  verschwindet und somit

$$\begin{aligned} \langle p|y \rangle &= p_n \frac{x_{k_m}}{c_m} + \langle p|x \rangle - \frac{x_{k_m}}{c_m} \sum_{l=1}^M p_{k_l} c_l = \langle p|x \rangle + \frac{x_{k_m}}{c_m} \left( p_n - \sum_{l=1}^M p_{k_l} (e_l)^t (A_X)^{-1} A e_n \right) \\ &= \langle p|x \rangle + \frac{x_{k_m}}{c_m} \left( p_n - \langle p_X | (A_X)^{-1} A e_n \rangle \right). \end{aligned} \quad (12.6)$$

Also wird  $n$  so gewählt, dass letztere Faktor möglichst negativ ist, d.h. gegeben ist durch

$$\eta_n = \min \{ \eta_l \mid l \in \{1, \dots, M\} \setminus X \}, \quad \eta_l = p_l - \langle p_X | (A_X)^{-1} A e_l \rangle. \quad (12.7)$$

Der folgende Satz bestätigt nun, dass die Auswahl von zunächst  $n$  gemäß (12.7) und dann  $m$  gemäß (12.5) tatsächlich einen funktionierenden Algorithmus darstellt.

**12.10 Satz** Sei  $n$  gemäß (12.7) gewählt, und  $c$  definiert durch (12.3).

- (i) Falls  $\eta_n \geq 0$ , so ist  $x$  eine minimierende Ecke.
- (ii) Falls  $\eta_n < 0$  und  $c < 0$ , so hat das Optimierungsproblem keine Lösung.
- (iii) Falls  $\eta_n < 0$  und nicht  $c \leq 0$ , sei  $m$  gemäß (12.5) bestimmt. Dann liefert (12.4) eine neue Ecke mit niedrigerem Kostenwert.

**Beweis.** (i) Sei  $y \in P$  beliebig. Dann gilt

$$x_X = (A_X)^{-1} A_X x_X = (A_X)^{-1} A x = (A_X)^{-1} b = (A_X)^{-1} A y = y_X + \sum_{l \notin X} y_l (A_X)^{-1} A e_l.$$

Einsetzen liefert

$$\begin{aligned} \langle p|x \rangle &= \langle p_X | x_X \rangle = \langle p_X | y_X \rangle + \sum_{l \notin X} y_l \langle p_X | (A_X)^{-1} A e_l \rangle \\ &= \langle p|y \rangle + \sum_{l \notin X} y_l \left( \langle p_X | (A_X)^{-1} A e_l \rangle - p_l \right) = \langle p|y \rangle - \sum_{l \notin X} y_l \eta_l \leq \langle p|y \rangle, \end{aligned}$$

wobei im letzten Schritt die Voraussetzung eingegangen ist. (ii) Wenn  $c < 0$ , so können wir anstelle von  $y$  nach (12.4) eine Einparameterfamilie  $y(\lambda)$  mit  $\lambda \geq 0$  konstruieren, die ganz in  $P$  liegt. Wähle  $m$  beliebig und ersetze  $\lambda = -\frac{x_{k_m}}{c_m} \geq 0$  in (12.4). Das liefert:

$$y(\lambda)_{k_l} = x_{k_l} - \lambda c_l \quad \text{für } l \neq m, \quad \text{und} \quad y(\lambda)_n = \lambda,$$

wobei wieder  $y(\lambda)_m = 0$  für  $m \notin X \cup \{n\} \setminus \{k_m\}$ . Offensichtlich gilt dann  $y(\lambda) \geq 0$  und außerdem

$$\begin{aligned}
 Ay(\lambda) &= \lambda Ae_n + \sum_{l=1}^M (x_{k_l} - \lambda c_l) Ae_{k_l} = A_X x_X + \lambda \left( Ae_n - \sum_{l=1}^M c_l Ae_{k_l} \right) \\
 &= Ax + \lambda \left( Ae_n - \sum_{l=1}^M \langle e_l | C_X e_n \rangle Ae_{k_l} \right) = b + \lambda \left( Ae_n - \sum_{l=1}^M Ae_{k_l} \langle e_l | (A_X)^{-1} Ae_n \rangle \right) \\
 &= b + \lambda (Ae_n - A_X (A_X)^{-1} Ae_n) = b.
 \end{aligned}$$

Somit ist  $y(\lambda) \in P$  für alle  $\lambda \geq 0$ . Für die Kosten gilt nach (12.6)

$$\langle p | y(\lambda) \rangle = \langle p | x \rangle + \lambda \eta_n,$$

was wegen  $\eta_n < 0$  beliebig klein wird. (iii) Dieser Punkt wurde schon vor der Formulierung des Satzes bewiesen.  $\square$

# A Äquivalenzrelationen

**A.1 Definition** Sei  $X$  eine Menge und  $X \times X = \{(x, y) \mid x, y \in X\}$  die Produktmenge von Paaren von Punkten in  $X$ . Eine Teilmenge  $\mathcal{R} \subset X \times X$  heißt eine Relation auf  $X$ . Eine Relation kann folgende Eigenschaften haben:

- (i)  $\mathcal{R}$  reflexiv  $\iff (x, x) \in \mathcal{R} \quad \forall x \in X$
- (ii)  $\mathcal{R}$  symmetrisch  $\iff ((x, y) \in \mathcal{R} \iff (y, x) \in \mathcal{R})$
- (iii)  $\mathcal{R}$  transitiv  $\iff ((x, y) \in \mathcal{R} \text{ und } (y, z) \in \mathcal{R} \implies (x, z) \in \mathcal{R})$

**A.2 Definition** Eine reflexive, symmetrische und transitive Relation  $\mathcal{R}$  heißt Äquivalenzrelation. Anstelle von  $(x, y) \in \mathcal{R}$  schreibt man dann oft auch  $x \sim y$  oder  $x \approx y$ .

**A.3 Beispiele** 1. Sei  $X$  eine Menge und  $f : X \rightarrow X$  eine Abbildung. Eine zugehörige Relation, auch genannt der Graph von  $f$ , ist

$$\mathcal{R} = \{(x, f(x)) \mid x \in X\} \subset X \times X.$$

Diese Relation  $\mathcal{R}$  ist reflexiv  $\iff f = \text{id}_X \iff \mathcal{R} = \{(x, x) \mid x \in X\}$  Diagonale.

2. Umgekehrt, sei gegeben eine Relation  $\mathcal{R}$  auf  $X$ . Dann gilt:

$$\begin{aligned} \mathcal{R} \text{ ist Graph einer Funktion} &\iff \forall x \in X \exists! y \in X \text{ mit } (x, y) \in \mathcal{R} \\ &\iff ((x, y) \in \mathcal{R} \text{ und } (x, z) \in \mathcal{R} \iff y = z) \end{aligned}$$

3. Sei  $p \in \mathbb{N}$ . Eine Äquivalenzrelation auf den ganzen Zahlen  $\mathbb{Z}$  ist definiert durch

$$\begin{aligned} n \sim m &\stackrel{\text{def.}}{\iff} n - m \text{ Vielfaches von } p \\ &\iff n - m \text{ durch } p \text{ teilbar} \\ &\iff n = m \pmod{p} \end{aligned}$$

◇

**Idee:** Fasse alle äquivalenten Elemente zu so genannten Äquivalenzklassen zusammen.

**A.4 Definition** Sei  $\sim$  eine Äquivalenzrelation auf  $X$ . Dann heißt  $[x] = \{y \in X \mid x \sim y\}$  die Äquivalenzklasse von  $x$ . Elemente aus  $[x] \subset X$  heißen Repräsentanten von  $[x]$ .

**A.5 Satz** Zwei Äquivalenzklassen  $[x], [x'] \subset X$  sind entweder gleich oder disjunkt.

**Beweis**

$$\begin{aligned} y \in [x] \cap [x'] &\iff y \sim x \text{ und } y \sim x' \\ &\stackrel{\text{Symmetrie}}{\iff} x \sim y \text{ und } y \sim x' \\ &\stackrel{\text{Transitivität}}{\iff} x \sim x', \text{ d.h. } [x] = [x'] \end{aligned}$$

Somit: nicht disjunkt  $\iff$  gleich (da die Umkehrung  $\Leftarrow$  trivial ist).

□

**A.6 Definition** Sei  $\sim$  eine Äquivalenzrelation auf  $X$ . Dann heißt die Menge der Äquivalenzklassen

$$X/\sim = \{[x] \mid x \in X\}$$

der Quotient von  $\sim$ .

## B Konvexitätsungleichungen

Dieser Anhang enthält einige Hilfsmittel der Analysis, die im Zusammenhang mit den  $p$ -Normen im Kapitel 11 verwandt werden.

**B.1 Definition** Sei  $I \subset \mathbb{R}$  ein Intervall. Eine Funktion  $f : I \rightarrow \mathbb{R}$  heißt konvex genau dann wenn für alle  $x, y \in I$  und  $\lambda \in [0, 1]$  gilt

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Eine Funktion  $f$  heißt konkav genau dann, wenn  $-f$  konvex ist.

**B.2 Satz** (Jensen's Ungleichung) Sei  $f : I \rightarrow \mathbb{R}$  konvex. Weiter seien  $x_1, \dots, x_N \in I$  und  $0 < \lambda_n < 1$  so, dass  $\sum_{n=1}^N \lambda_n = 1$ . Dann gilt

$$f\left(\sum_{n=1}^N \lambda_n x_n\right) \leq \sum_{n=1}^N \lambda_n f(x_n).$$

**Beweis** durch Induktion über  $N$ . Die Aussage ist die für  $N = 2$  genau die Definition der Konvexität. Somit muss lediglich der Schritt von  $N$  nach  $N + 1$  nachgewiesen werden. Setze  $\lambda = \sum_{n=1}^N \lambda_n$  und  $x = \frac{1}{\lambda} \sum_{n=1}^N \lambda_n x_n$ . Dann  $1 - \lambda = \lambda_{N+1}$  und somit

$$\begin{aligned} f(\lambda x + (1 - \lambda)x_{N+1}) &\leq \lambda f(x) + (1 - \lambda)f(x_{N+1}) \\ &= \lambda f\left(\sum_{n=1}^N \frac{\lambda_n}{\lambda} x_n\right) + \lambda_{N+1}f(x_{N+1}) \\ &\leq \lambda \sum_{n=1}^N \frac{\lambda_n}{\lambda} f(x_n) + \lambda_{N+1}f(x_{N+1}) \\ &= \sum_{n=1}^{N+1} \lambda_n f(x_n), \end{aligned}$$

wobei im vorletzten Schritt die Induktionsvoraussetzung verwandt wurde. □

**B.3 Satz** Sei  $1 \leq p < \infty$  und  $q \in \mathbb{R}$ , so dass  $\frac{1}{p} + \frac{1}{q} = 1$ . Setze

$$\|v\|_p = \left(\sum_{n=1}^N |v_n|^p\right)^{\frac{1}{p}}, \quad v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{K}^N,$$

und  $\|v\|_\infty = \max\{|v_1|, \dots, |v_N|\}$ . Dann gilt:

(i) (Hölder-Ungleichung)  $|\langle v | w \rangle| \leq \|v\|_p \|w\|_q$

(ii) (Minkowski-Ungleichung)  $\|v + w\|_p \leq \|v\|_p + \|w\|_p$

**Beweis.** (i) Der Fall  $p = 1, q = \infty$  folgt direkt aus

$$|\langle v | w \rangle| \leq \sum_{n=1}^N |v_n| |w_n| \leq \left(\max_{n=1, \dots, N} |w_n|\right) \sum_{n=1}^N |v_n| = \|w\|_\infty \|v\|_1.$$

**Behauptung** Für  $a, b > 0$  gilt  $ab \leq \frac{1}{p} a^p + \frac{1}{q} b^q$  (Young'sche Ungleichung).

**Begründung** Die Funktion  $x \in \mathbb{R}_{>0} \mapsto \ln(x)$  ist konkav, also für  $x, y > 0$

$$\frac{1}{p} \ln(x) + \frac{1}{q} \ln(y) \leq \ln\left(\frac{1}{p} x + \frac{1}{q} y\right).$$

Da exp monoton ist, folgt nach Exponentieren

$$x^{\frac{1}{p}} y^{\frac{1}{q}} \leq \frac{1}{p} x + \frac{1}{q} y.$$

Setze  $x^{\frac{1}{p}} = a$ ,  $y^{\frac{1}{q}} = b$ , dann ist die Behauptung gezeigt. ◇

Nun sei  $\|v\|_p \neq 0$  und  $\|w\|_q \neq 0$  (sonst ist die Aussage trivial). Es folgt:

$$\frac{\langle v | w \rangle}{\|v\|_p \|w\|_q} \leq \sum_{n=1}^N \frac{|v_n| |w_n|}{\|v\|_p \|w\|_q} \leq \sum_{n=1}^N \left( \frac{1}{p} \frac{|v_n|^p}{\|v\|_p^p} + \frac{1}{q} \frac{|w_n|^q}{\|w\|_q^q} \right) = \frac{1}{p} + \frac{1}{q} = 1.$$

Dies zeigt (i). Für  $p = 1$  folgt die Minkowski-Ungleichung direkt aus der Dreiecksungleichung. Seien also  $1 < p$  und  $q < \infty$  mit

$$\frac{1}{p} + \frac{1}{q} = 1 \iff q + p = pq \iff p = (p-1)q \iff p - \frac{p}{q} = 1.$$

Dann

$$\begin{aligned} \|v + w\|_p^p &= \sum_{n=1}^N |v_n + w_n| |v_n + w_n|^{p-1} \\ &\leq \left( \sum_{n=1}^N |v_n| |v_n + w_n|^{p-1} \right) + \left( \sum_{n=1}^N |w_n| |v_n + w_n|^{p-1} \right) \\ &\stackrel{\text{Hölder}}{\leq} (\|v\|_p + \|w\|_p) \left( \sum_{n=1}^N |v_n + w_n|^{(p-1)q} \right)^{\frac{1}{q}} \\ &= (\|v\|_p + \|w\|_p) \|v + w\|_p^{\frac{p}{q}} \quad \text{da } (p-1)q = p. \end{aligned}$$

Für  $\|v + w\|_p \neq 0$  (ansonsten ist die Aussage trivial) verwenden wir nun  $(\|v + w\|_p)^{p-\frac{p}{q}} = \|v + w\|_p$ . □

## Literatur

[ABHKS] T. Arens, R. Busam, F. Hetlich, C. Karpfinger, H. Stachel, *Grundwissen Mathematikstudium*, (Springer, 2013).

[Bri] Brieskorn, *Lineare Algebra und Analytische Geometrie*, (Vieweg, Braunschweig, 1983).

[Fis] G. Fischer, *Lineare Algebra*, (Vieweg, 2010).

[Gre] W. Greub, *Lineare Algebra*, (Springer, 1976).

[HJ] R. A. Horn, C. R. Johnson, *Matrix Analysis*, Second Edition (Cambridge Univ. Press, 2012).

[HW] B. Huppert, W. Willems, *Lineare Algebra*, (Teubner, Wiesbaden, 2006).

[Kna] A. Knauf, *Lineare Algebra und analytische Geometrie I und II*, Skript 2000 und 2001.

[Rom] S. Roman, *Advanced Linear Algebra*, (Springer, 1992).

[SW] R. Schaback, H. Wendland, *Numerische Mathematik*, (Springer, 2005).